Investigation of the Epstein-Barr virus' Latent to Lytic Cycle Switch in Epithelial Cells Suggests an Interplay Between Host Transcription Co-Factors and the Viral Early Lytic Cycle Protein Zta.

By Anja Katharina Godfrey

A thesis submitted for the degree of

Doctor of Philosophy

University of Sussex

September 2018

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:.....

Acknowledgements

First and foremost I would like to thank my supervisor, Professor Alison Sinclair. Throughout the course of this PhD I constantly felt fortunate to be her student. I am grateful for the space she has given me to learn, while always being available when support was required. Her patience, strength, knowledge, tact, intelligence, understanding, resilience and dedication make her an extraordinary supervisor. I have learned a lot from her.

I would also like to thank my co-supervisor, Professor Michelle West, without whom I would not have had the opportunity to do this PhD. I am also grateful for her feedback, passion and advice.

Thank you to past and present members of the Sinclair lab group, Dr Christopher Traylen, Dr Ijiel Barak Naranjo Perez Fernandez, Dr Rajaei Almohammed and Faisal Alzahrani, for their help and friendship. Special thanks to Dr Kay Osborne and Yaqi Zhou for their continuous patience, help, and for sharing their skills. Thank you also to all past and present members of the West lab for their help, friendship and advice, especially Dr David Wood and Dr Michael McClellan.

Finally, thank you to my family. To my parents, Birgit and Günter Jahnkow: Thank you for your love, care and boundless support. Thank you for working very hard to provide your daughters with a better, more privileged life than you. I am grateful that you raised me with the knowledge that I am responsible for my own life and with the firm believe that I can do anything I set my mind to. To my sister, Katrin Jahnkow: Thank you for your love and your belief in me. I am very grateful to have you on my side. To my parents-in-law, Lallie and Hugh Godfrey: Thank you for your care and support. To my clever, funny, caring and strong husband, Dr Sam Godfrey: Thank you. Thank you for your infinite love and unshakable belief in me, for being my rock and for making me happy. Everything I achieve, I achieve as part of the team that is you and I. To my inquisitive, compassionate, determined and clever daughter, Matilda Godfrey: During this PhD, you made us happier than I could have ever imagined. I am especially grateful to you for making me realise my strength and for continuously reminding me what is most important in life. Without you this PhD would not have been as enjoyable.

UNIVERSITY OF SUSSEX

ANJA KATHARINA GODFREY

PhD BIOCHEMISTRY

INVESTIGATION OF THE EPSTEIN-BARR VIRUS' LATENT TO LYTIC CYCLE SWITCH IN EPITHELIAL CELLS SUGGESTS AN INTERPLAY BETWEEN HOST TRANSCRIPTION CO-FACTORS AND THE VIRAL EARLY LYTIC CYCLE PROTEIN ZTA

SUMMARY

The Epstein-Barr virus (EBV) is one of the oldest and most successful human viruses, infecting over 95% of the human population. Exposure usually occurs during early childhood and leads to mild, or no, acute symptoms. After primary infection, the virus enters naive B cells, causing differentiation and proliferation, finally leading to lifelong persistence of the virus in these cells (latency). Infectious virions are produced during the lytic phase of the EBV life cycle, leading to spread of the virus via oral secretions. There is a strong link between EBV infection and over a dozen malignant diseases. A large amount of our EBV knowledge has been obtained in B cells. However, this virus also infects epithelial cells and is associated with a range of epithelial malignancies. This thesis aims to elucidate some of the ambiguity surrounding EBV lytic cycle in epithelial cells and to extend the knowledge surrounding the switch from latency to lytic cycle. For this purpose EBV was induced into lytic cycle, in the epithelial cell line HONE1-EBV, with the histone deacetylase inhibitor SAHA, which is used in the treatment of non-Hodgkin lymphoma and is in Phase II clinical trials for the treatment of nasopharyngeal carcinoma. The binding of the early lytic cycle protein, Zta, to the epithelial host cell genome was investigated using Chromatin Immunoprecipitation followed by high-throughput sequencing. This led to the identification of several host genes regulated by the lytic virus. In addition, results previously obtained in B cells and other public ChIP sequencing data were included in the data analysis and revealed a potentially significant role of the host transcription co-factor and proto-oncogene candidate BCL3. Further experiments led to the proposal of a model for the regulation of Zta gene activation involving Zta, BCL3 and TORC.

Table of Contents

List of Figures

List of Supplementary Figures

List of Tables

Table of Abbreviations

Abbreviation	**Elaboration**
aa	Amino Acid
AP	Activator Protein
BALF	BamHI A left reading frame transcript
*Bam*HI	Bacillus Amyloliquefaciens Restriction Enzyme H1
BART	BamHI-A rightward transcripts
BCA	Bicinchoninic Acid

BCL2	B cell Lymphoma 2 Protein
BCL3	B cell Lymphoma 3 Protein
BL	B Lymphocyte
BLLF	BamHI L left reading frame transcript
BMRF	BamHI M right reading frame transcript
bp	Base Pairs
BRLF	EcoRI left reading frame transcript
BZLF	BamHI Z left reading frame transcript
C	Celcius
CBP	CREB Binding Protein
CD	Cluster of Differentiation
cDNA	complementary DNA
ChIP	Chromatin Immunoprecipitation
ChIP-Seq	Chromatin Immunoprecipitation Coupled to Sequencing
Cp	Latency Promoter within BamHI C Digestion Fragment
CpG	Cytosine-phosphate-Guanine
CREB	cAMP Response Element Binding Protein
CTAR	Carboxy-terminal Activating Region
DNA	Deoxyribonucleic Acid
EBER	Epstein-Barr Virus Encoded Ribonucleic Acid
EBNA	Epstein Barr Nuclear Antigen
EBV	Epstein-Barr Virus
EGFR	Epidermal Growth Factor Receptor

ENCODE	Encyclopaedia of DNA Elements
FACS	Fluorescence Activated Cell Sorting
FBS	Foetal Bovine Serum
FTP	File Transfer Protocol
GFP	Green Fluorescent Protein
gH	Envelope Glycoprotein H
gL	Envelope Glycoprotein L
GO	Gene Ontology
gp	Glycoprotein
GREAT	Genomic Regions Enrichment of Annotations Tool
HEK	Human Embryonic Kidney
HeLa	Henrietta Lacks
HHV	Human Herpes Virus
HIV	Human Immunodeficiency Virus
HLA	Human Leukocyte Antigen
hrs	Hours
HTLV-1	Human T-cell Leukaemia Virus Type 1
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IL	Interleukin
IκB	Inhibitor of κB
kDa	Kilo Dalton
kb	Kilobases

LCL	Lymphoblastoid Cell Line
LMP	Latent Membrane Protein
M	Molar
m	Milli
MACS	Model Based Analysis of ChIP Sequencing
MEME-ChIP	Motif Analysis of Large Nucleotide Datasets
min	Minutes
miRNA	Micro RNA
mRNA	Messenger RNA
NCBI	National Centre for Biotechnology Information
NFκB	Nuclear factor kappa enhancer of activated B cells
n	Nano
nm	Nanometers
NPC	Nasopharyngeal Carcinoma
OriLyt	Origin of Lytic Replication
OriP	Origin of Plasmid Replication
p (significance)	Probability
p (e.g chromosome region 3p)	p-arm (short arm) of human chromosome (e.g. 3)
PANTHER	Protein Analysis through Evolutionary Relationships
PBS	Phosphate Buffered Saline
pCREB	Phosphorylated CREB
qPCR	Quantitative Polymerase Chain Reaction

RNA	Ribonucleic Acid
SAHA	Suberoylanilide Hydroxamic Acid
SD	Standard Deviation
sec	Seconds
siRNA	small interfering RNA
STAT	Signal Transducer and Activator of Transcription
STR	Short Tandem Repeat
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
TLR	Toll-like Receptor
TNF	Tumour Necrosis Factor
TORC	Transducer of Regulated CREB activity
TR	Terminal Repeat
UCSC	University of California, Santa Cruz
UTP	Uridintriphosphat
VCA	Viral Capsid Antigen
Wp	Latency Promoter within BamH1 W Digestion Fragment
ZKO	Zta Knock-Out
Zp	Promoter of BZLF1
ZRE	Zta Response Element
µ	Micro

Chapter 1: Introduction

1.1 General Overview

1.1.1 Brief Introduction to Viruses

Any virus on its own has no metabolism and cannot replicate; it is an acellular, inert complex of molecules. But when viruses enter cells, they can be replicated, which causes harm to the host. Viruses can vary greatly in shape, size and content. They generally consist of some nucleic acid - single or double strand ribonucleic acid (RNA) or deoxyribonucleic acid (DNA), surrounded by a protective protein shell, the nucleocapsid. Many viruses also carry replicase enzymes and some viruses have an additional membrane [1]. The number of genes encoded by different types of virus can range from three to several hundred. Similarly their size can vary greatly, ranging from 10nm-300nm [2]. Most eukaryotic viruses can only infect certain cell types, depending on specific host cell membrane surface receptors which can be bound by the virus and lead to cell entry via endocytosis or direct fusion with the membrane. Cell types that lack specific receptors are much less prone to viral infection. Bacteriophages, which are often used in gene cloning, attach to the bacterial cell wall using their tail fibre through which they then inject their DNA into the cell [1, 3].

Once inside the cell, some viral genomes can be integrated into the host cell chromosome as an additional set of genes which are only replicated as the cell divides. This is known as latency. Most of the symptoms attributed to a viral infection are caused by the active replication of the virus independent of the host cell chromosome. This process leads to the destruction of the host cell and is known as lytic cycle, which will be the focus of this thesis. Following viral

replication, newly formed viruses are released. This process may kill the infected host cell through lysis. In other cases the viral release causes no immediate damage but leads to the embedding of some viral proteins in the cell membrane, which allow the immune system to recognize infected cells. If the viral antigen is associated with an MHC class I molecule, natural killer cells bind to the foreign viral proteins and cause membrane trauma leading to cell death. In cases where the immune system does not recognize an infected cell, viruses can cause apoptosis by producing deficiency injury in infected cells. During active viral replication, essential metabolites are in such heavy demand that the cell cannot survive [1].

Infectious diseases are the second most common cause of death worldwide [4]. While bacterial infections can mostly be controlled by antibiotics, finding drugs that effectively treat viral infections is very challenging. Viruses use the host cell's resources and metabolism, so effective antiviral drugs often lead to toxicity for the host. As is the case in chemotherapeutic cancer treatment, some licensed antiviral drugs are more toxic to the virus than to the host (e.g. nucleoside analogues). However, the development of drug resistant viruses is common in this type of treatment. Some agents produced by the host itself (i.e. Interferons), which can inhibit cell protein synthesis, are able to target viruses by preventing viral spread to further cells. However, only small amounts of interferons are produced in response to highly virulent viruses. In addition, interferons are host-specific, so interferons produced in one species cannot be used to treat viral infection in another [5].

1.1.2 Herpesviruses

Herpesviruses make up a large family of viruses that cause many different diseases including cold sores, chicken pox and infectious mononucleosis. Some herpesviruses can remain latent in host cell for extended periods of time (years) and only become active in response to stress conditions. In addition herpesviruses, like the Epstein-Barr virus, can cause cancer. The structure of herpesviruses is quite complex (Figure 1) [2, 5].

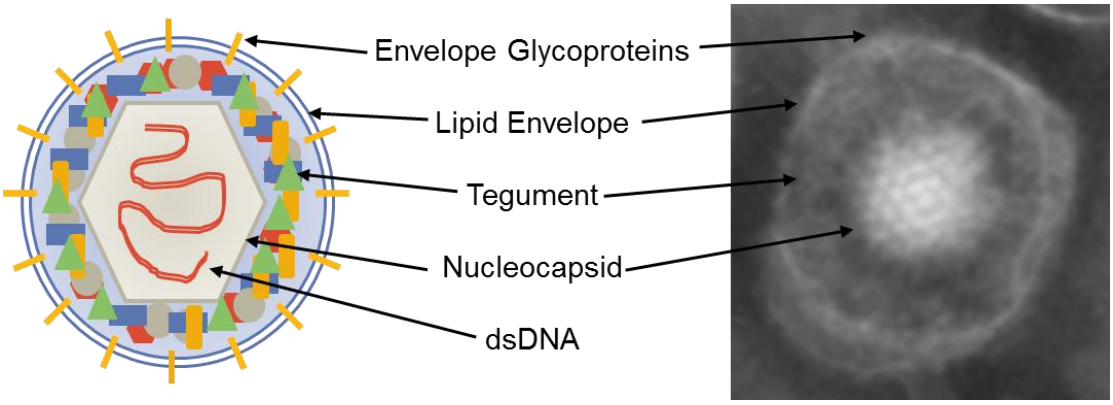


Figure 1: Schematic image and electronmicrograph of herpesvirus structure. Figure shows an enveloped polyhedral capsids with linear double-stranded DNA. Electronmicrograph from Goldsmith et al., 2009 [6].

Herpes viruses infect host cells by attaching to specific cell receptors, fusing the cell membrane with the viral envelope. This allows the nucleocapsid to be released into the cell before it is transported to the nucleus, where the viral DNA is uncoated, circularizes and latency is established [5]. When the virus enters lytic cycle, the genome is replicated, viral genes are expressed and the virus is

assembled in the nucleus. The viral envelop is obtained from the host cell's nuclear membrane and the fully-formed virus exits via exocytosis or lysis.

Herpesviruses have very high infection rates and are the most prevalent type of DNA virus [2]. Almost all animal species are natural hosts to at least one type of herpesvirus. Overall, more than 200 herpesviruses have been identified which have particular specificity for certain host species and are classified into three subfamilies: α, β and γ. In addition to this classification, a numbering system has been assigned to the eight known human herpesviruses (HHV) with numbers corresponding to the order of discovery (except HHV-7). [1, 2]

1.1.3 The Epstein - Barr virus

The Epstein-Barr virus (EBV, HHV-4) is the only γ-1 herpesvirus (or lymphocryptovirus) that is known to infect humans and humans are its only natural host [11]. EBV was first discovered in 1964 by Anthony Epstein, Yvonne Barr, and Burt Achong in cultured tumour cells from a Burkitt Lymphoma biopsy [6, 7]. Since then it has become clear that EBV is one of the oldest, most common human viruses which has infected mankind ever since we parted from our primate ancestors [7]. EBV is highly prevalent all over the world, even in remote locations. More than 95% of the world's population are infected by this virus, which makes EBV one of the most successful human pathogens [8, 9].

In 2012, 14 million people were diagnosed with cancer, of which 2.2 million cases (15.4%) were directly linked to infections [10]. EBV is one of the most common causes of infection associated cancer, with an estimated 200 000 new cases annually [11]. While EBV infection is benign and asymptomatic in the majority of

cases, there is a strong link between EBV infection and over a dozen malignant diseases in both immunocompromised and immunocompetent patients. In total, 1.5% of cancer cases are associated with EBV [10]. The four most common malignancies connected to EBV infection are Hodgkin's lymphoma, Burkitt lymphoma, gastric carcinoma and non-glandular nasopharyngeal carcinoma (NPC), which will be discussed in more detail further on in this introduction. EBV causes approximately 10% of gastric carcinomas, which are the second most common cause of cancer deaths [12]. Many factors point to the causative role of EBV in malignant transformation. These include viral oncogenes and the presence of EBV in every cancerous cell in EBV associated tumours. In immunocompetent hosts, cancer detection occurs several years after primary infection, suggesting that EBV infection is only one factor in the complex, multistep process of malignant transformation [6]. Cofactors which are able to initiate the lytic stage of the viral life cycle, thus reactivating viral replication, are most likely required for EBV associated oncogenesis [11]. The International Agency for Research on Cancer classifies EBV as a Group 1 carcinogen [12].

In addition to the direct link between EBV and cancer, increasing evidence points to the implication of EBV in the pathogenesis of Multiple Sclerosis, a chronic inflammatory demyelination of the central nervous system which leads to progressive disability in affected individuals [13].

Evolution of EBV, and other herpes viruses, has been driven by survival rather than the ability to cause illness. Persistence and reactivation have become key to the success of this group of pathogens [6]. Persistence of EBV can be detected as a lytic infection in the oral mucosal epithelium in immunocompromised people, which leads to the spread of the virus via oral secretion. EBV can also be

transmitted through blood transfusions. EBV exposure often takes place during early childhood and leads to mild, or no, acute symptoms. It has been suggested that persistent maternal antibodies might play a protective role in containing the infection at this stage [14]. The primary infection stimulates a specific immune response, but this is unable to rid the body completely of the virus [15]. When exposure is delayed to a later point in life, which is more common in developed countries, EBV infection leads to infectious mononucleosis in about 50% of primary infection cases. Infectious mononucleosis is a self-limiting lymphoproliferative disease characterised by fever, pharyngitis and general lymphadenopathy [12, 16]. After primary infection, the virus infects naive B cells and expresses all of the EBV genes associated with latency including 10 proteins and two small RNAs (latency type III), which drive B cell transformation and proliferation but also initiate an immune response which leads to the elimination of infected B cells by EBV-specific T cells. By downregulating its immunogenic proteins, the virus is able to survive within the host. Infected B cells enter lymph nodes where they are able to proliferate while only expressing three viral proteins (latency type II). When the infected B cells exit the lymph nodes as non-dividing memory B cells, all viral proteins are downregulated (latency type 0), making the virus undetectable to the immune system. When the infected B cells divide, the virus expresses one protein (latency type I) to ensure that the viral genome divides with the host cell genome [17]. A model of the EBV life cycle is visualised in Figure 2.

Even though EBV is probably the most potent transforming human virus in cell culture [12], it persists in the nucleus of memory B cells for the duration of the host's life, often asymptomatically as the viral genome. During this time, a

relatively small amount of active viral replication occurs which causes transmission of the virus via oral secretions [6, 18]. All γ-1 herpesvirus have a distinct set of genes which enables them to persist in B cells. EBV is also able to infect other cell types, including mesenchymal cells, T cells, and epithelial cells which will be particularly relevant for this thesis.

EBV has a linear 172kb double stranded DNA genome (Figure3), which encodes >85 genes and is wrapped around an O-shaped protein core. When the virus infects a cell, the ends of the linear genome join together to form circular, episomal DNA. The genome contains long and short unique sequence domains; these sections are responsible for almost all of its coding capacity. Other parts of the EBV genome include internal and terminal tandem and reiterated direct repeats. The number of joined terminal repeats differs between virions due to the variable cleavage of the genome during the replication process. This makes it possible to distinguish clonal infection events. The number of terminal repeats stays constant in dividing, latently infected cells [6, 7].

There are two major types of EBV, namely EBV-1 and EBV-2 (or type A and B), which contain substantial differences in DNA sequence, particularly in their nuclear antigen genes (EBNA2, EBNA3A, EBNA3B and EBNA 3C). In addition, there are different viral strains within each type based on heterogeneity of certain genomic regions. EBV strains vary between different geographical regions, but there can also be a range of different strains in the same area [19, 20].

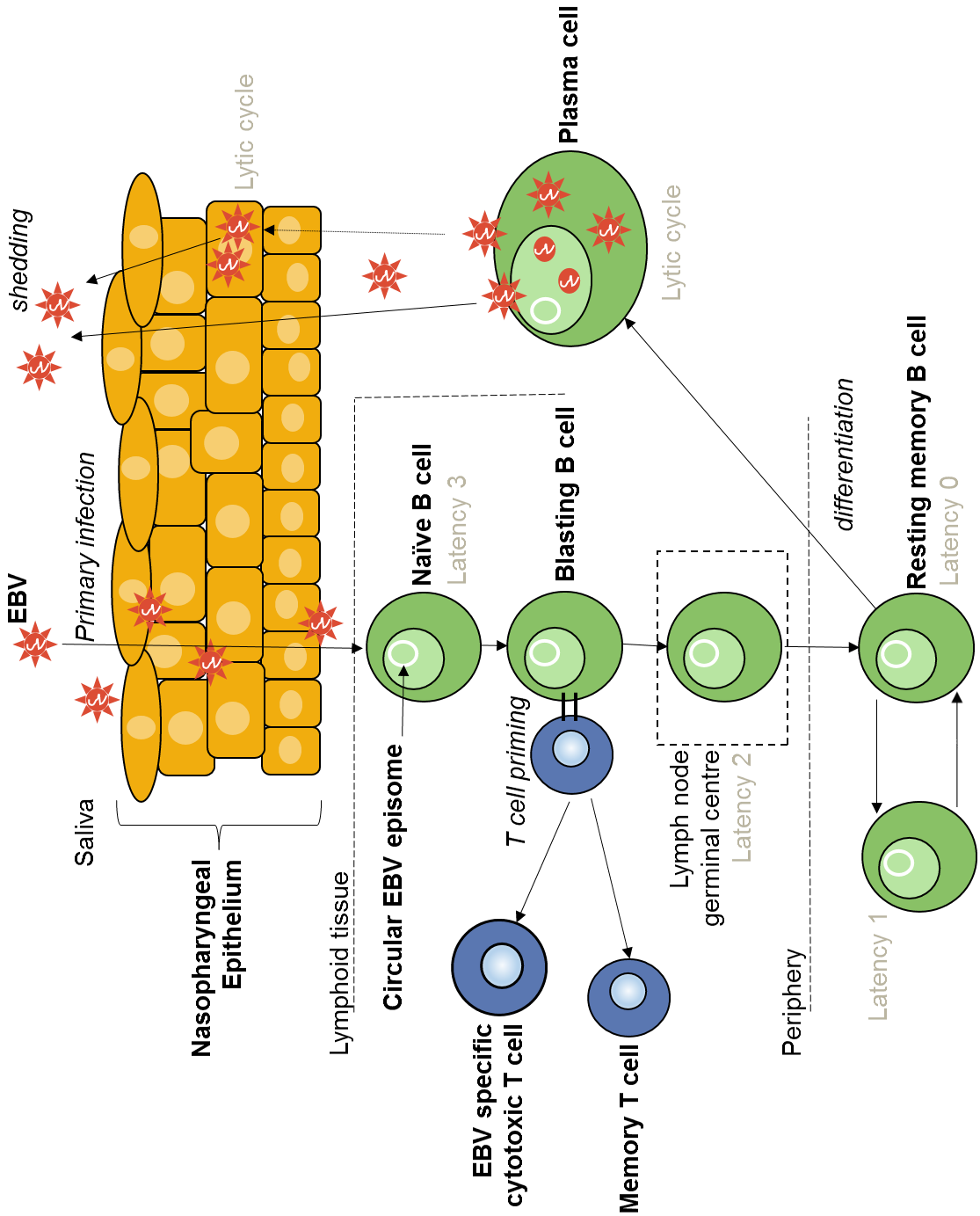


Figure 2: Model of EBV life cycle.

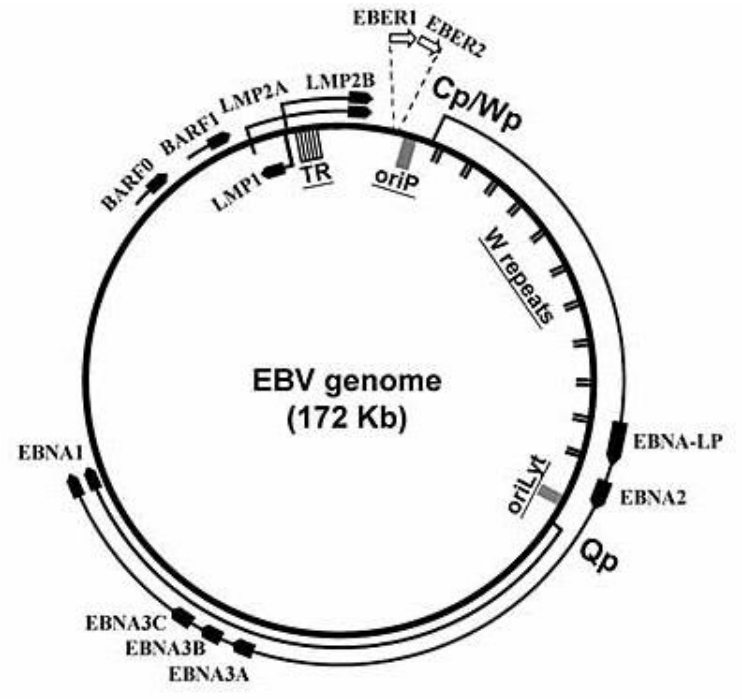


Figure 3: Map of circular, double-stranded EBV episome. Figure shows latent genes. Modified from Tao *et al.*, 2006 [21]. oriP: origin of plasmid replication, oriLyt: origin of lytic replication, TR: terminal repeat. The locations and transcription of viral latent genes are indicated. The outer arrows show the large EBV transcript in latency 3, during which all the EBNAs are transcribed from either the Cp or Wp promoter. The EBNAs arise from the same transcript that is differentially spliced. The inner arrow shows EBNA1 originating from the Qp promoter which is transcribed in latency 1 and 2.

1.2 EBV in B cells

1.2.1 Entry into B cells

EBV enters B cells by interacting with the complement receptor CD21 on the cell's surface via the viral envelop glycoprotein gp350/220 and binding of a second glycoprotein, gp42, to human leukocyte antigen class II which acts as a co-

receptor. When the viral molecules interact with the B cell receptors, CD21 aggregates within the membrane and EBV is endocytosed in cytoplasmic vesicles [6, 15]. Fusion in this low pH compartment occurs via three proteins which are part of the core fusion machinery [22], common to all herpesviruses. The core fusion machinery is activated by the interaction of gp42 with HLA class II, which leads to the widening of a hydrophobic pocket on the gp42 protein [23], and is made up of homodimers of glycoprotein gB and a heterodimer of glycoproteins gH and gL. Binding of glycoprotein gp42 to gH allows the formation of the gHgLgp42 complex. This complex leads to the activation of gB which is thought to insert a fusion loop into the cell membrane and undergo a conformational change which leads to the fusion of the viral and cellular membranes. [24]

1.2.2 Latency in B cells

The unique ability of EBV to convert resting B cells into permanently latently infected lymphoblastoid cell lines (LCLs) has provided researchers with an invaluable, yet incomplete, *in vitro* model of the lymphomagenic potential of EBV [15]. A large proportion of our knowledge about EBV relates to B lymphocytes. When EBV is in the latent phase of its life cycle, no infectious viruses are produced and only a highly restricted set of latent genes are expressed, which contribute to the ability of the virus to evade immune detection. The expressed genes disrupt normal B cell growth which leads to virus-induced immortalisation and the establishment of LCLs [6].

The EBV genome is maintained as a circular episome during latency with every LCL carrying multiple extrachromosomal copies. A limited set of viral genes is

constitutively expressed during latency, comprising of six viral nuclear antigens, EBNAs 1, 2, 3A, 3B, 3C and –LP, and three latent membrane proteins, LMPs 1, 2A and 2B. Transcripts from the *Bam*HIA region and non-coding EBER1 and EBER2 can also be detected [15]. Four distinct latency programs have been identified, including the above mentioned latency type 0 (no latent gene products can be detected), latency type 1 (EBER and BART transcripts are present and EBNA1 is expressed), latency type 2 (EBER and BART transcripts, EBNA1, LMP1 and LMP2) and latency type 3 (all latent gene product are expressed, including EBNAs 1, 2, 3A, 3B, 3C and –LP, all three latent membrane proteins LMP1, LMP2A and LMP2B, as well as the EBER and BART RNAs) [25]. In addition, some subsets of BL tumours express EBNA1, -3A, -3B, and -3C and a truncated EBNA-LP without LMPs or EBNA2. This is known as Wp restricted latency as transcription is limited to the Wp promoter [26]. Table 1 summarises the EBV latency types.

Latency Type	**Viral Gene Expression**	**Occurrence**
0	none	Non-dividing, infected memory B cells
1	EBER and BART miRNAs EBNA1	Dividing, infected B cells, Burkitt's Lymphoma, Primary Effusion Lymphoma
2	EBER and BART miRNA, EBNA1, LMP1 and LMP2	Recently infected B cells in lymph nodes, NPC, Gastric carcinoma,

		Hodgkin's lymphoma, T/NK cell lymphoma
3	EBNAs 1, 2, 3A, 3B, 3C and –LP, LMP1, LMP2A and LMP2B, EBER and BART miRNAs	Primary infection, AIDS_NHL (large cell), PT-LPD
Wp restricted	EBNA1, -3A, -3B, and -3C and a truncated EBNA-LP	Subsets of BL tumours

Table 1: EBV latency types. Table includes genes expressed in each latency type and when you would expect to see that type of latency.

During latency, the viral genome is replicated by host cell DNA polymerase and is equally distributed to daughter cells during mitosis. Latent infection of B cells is essential for the life-long persistence of the virus. It has been shown that patients who lack B cells (X-linked agammaglobulinemia) do not show evidence of persistent EBV infection [27].

The main oncogenic EBV protein is the latent membrane protein LMP1 which is essential for tumorigenic B cell transformation [28, 29]. LCLs express anti-apoptotic proteins like BCL2 and A20, and high levels of B cell activation markers including CD23, CD30, CD39 and CD70 as well as the cell adhesion molecules CD11a, CD58 and CD54 which are normally expressed at very low levels in resting B cells until they are activated by antigenic or mitogenic stimuli. This indicates that EBV-induced immortalization can be achieved through the constitutive activation of the same pathways that drive physiological B-cell proliferation [15, 30].

1.2.3 Lytic cycle in B cells

Lytic replication is initiated in B cells when the latent virus is reactivated following physiological stimuli or chemical assault, such as foreign antigens, immunodeficiency, malaria, protein kinase C agonists, DNA methyltransferase inhibitors, histone deacetylase inhibitors or anti-IgG [12, 18]. Key early lytic promoters and the origins of lytic replication are in a repressive heterochromatin-like environment during latency which has to be overcome to achieve viral replication [18]. During the lytic cycle, the EBV genome goes through substantial epigenetic modifications including chromatin remodelling and DNA methylation [18]. The earliest indicators of the lytic cycle are the proteins Zta and Rta which are key mediators in the switch from latency to lytic cycle as they act as powerful transactivators of later lytic cycle genes and lead to the expression of the viral DNA polymerase [31]. During the lytic phase of the EBV life cycle, infectious viruses are produced. At this stage approximately 80 viral proteins are expressed which include DNA replication factors, transcriptional activators and structural proteins. After replication, the linear genome is cleaved at the terminal repeats and packaged into viral capsids. Finally, the virus is enveloped and released from the cell [6].

1.3 EBV in Epithelial cells

1.3.1 Entry into epithelial cells

Entry of EBV into epithelial cells occurs through separate, as yet poorly defined pathways, that seem to require EBV gH and gL glycoproteins but not gp42 as epithelial cells lack HLA class II expression [32, 33]. EBV infection of epithelial cells is much less efficient than B cell infection and may involve a direct transfer

infection from EBV-coated B cells to epithelial cells. The surface molecule α5β1 integrin may contribute to EBV infection in this cell type [6]. It has been shown that tonsillar and adenoid epithelial cells express CD21 mRNA, which may suggest that this protein could play a similar role to infection in B cells. [34, 35] A viral membrane protein which is encoded by the bmrf2 open reading frame can bind β1 integrin and may also be involved in attachment. [36, 37] gH is likely involved in the fusion of the viral and cellular membrane. Fusion does not appear to require endocytosis, and the resulting low pH, as is the case in B cells. Due to the lack of HLA class II, gp42 does not seem to interact with the gHgL complex. Recent studies have shown that the ephrin receptor tyrosine kinase A2 specifically binds to gHgL and is involved in the fusion of EBV to the cell membrane in epithelial cells. [38]

In vitro infection of epithelial cells does not activate the full growth-transforming programme of EBV and rarely achieves full lytic replication. In addition the EBV genome is lost during prolonged culture. Hence LCLs have remained the dominant *in vitro* model for studying EBV infection [15, 39].

1.3.2 Micro RNAs

In recent years there has been an increasing focus on the presence of microRNAs in EBV infection which seem to play a role in the transcription of the cellular and viral genome. EBV was the first virus shown to encode microRNAs [40]. It expresses at least 22 distinct microRNAs [41]. 14 microRNAs are located in the introns of the viral BART gene and are expressed at high levels in latently infected, differentiated epithelial cells while they are almost undetectable in B cells. More research is required to characterise their concrete function as they

are likely to play important roles in the viral life cycle and may contribute to the infection of different human tissues [25, 42].

1.3.3 Lytic Cycle in Epithelial Cells

It has been suggested that lytic replication of EBV in epithelial cells is rare, usually abortive and mainly occurs in small subsets of malignant cells [39]. The early lytic cycle protein Zta has occasionally been detected in NPC samples, in a minority of tumour cells [43]. Consistent expression of other lytic proteins and linear forms of DNA is absent within tumour cells, suggesting that NPCs do not fully permit viral replication [44, 45]. However, EBV lytic replication is considered an oncogenic risk factor as a rise in IgA antibodies to lytic viral proteins can frequently be detected in NPC patients prior to treatment (e.g. VCA and Zta). The appearance of these antibodies precedes NPC development by several years and has been sown to correlate to tumour burden [43, 46, 47]. While it has not been possible to definitively demonstrate the presence of lytically infected epithelial cells in healthy EBV carriers, the virus can be detected in throat washings. These findings indicate that reactivation of lytic cycle is a feature of EBV infection in healthy carriers [39], though it is uncertain whether these virions are produced by B cells or epithelial cells. Cell-free DNA can be detected in the serum of NPC patients, but it is not clear whether this represents encapsidated virions or naked DNA from dead host cells. It appears resistant to DNAse digestion and the NPC tumour burden is relatively small in comparison to the serum DNA levels, suggesting that it may arise from virions rather than dying NPC cells. This could point to some unrecognised role of lytic EBV replication in the pathogenesis of NPC [6]. In addition, studies have shown a strong association

between the levels of circulating tumour-derived EBV DNA in the blood of patients with NPC and overall survival rates, with low post-treatment levels leading to a good prognosis for the patient. It is thought to accurately reflect the residual tumour load [48].

EBV lytic replication in epithelial cells is indisputably observed in patients infected by the human immunodeficiency virus (HIV) with oral hairy leukoplakia, which is a non-malignant epithelial lesion located on the tongue where EBV replicates in the upper-layer of skin cells which are undergoing terminal differentiation, apoptosis and desquamation. This suggests that EBV replication is dependent on the differentiation state of epithelial cells [39]. When epithelial cells are grown in organotypic cultures (“rafts”), allowing the cells to differentiate and stratify as they do *in vivo*, EBV infection results in spontaneous production of large numbers of virions, specifically in the differentiating suprabasal cells. This may lead to the increase of the virus pool and increasing efficiency of transmission *in vivo*. [49]

1.3.4 EBV contribution to Nasopharyngeal Carcinoma

Certain types of EBV latency have been linked to specific cancers. Nasopharyngeal carcinoma (and gastric carcinoma) arises from latency 2 [50]. A diagram summarising the transformation of normal epithelium to NPC is shown in Figure 4.

Epithelial cancers associated with EBV are relatively rare globally. The type II non-keratinizing nasopharyngeal carcinoma shows the most consistent worldwide association with EBV with a high incidence in Southeast Asian, Alaskan Inuit, and North African populations, reaching a peak incidence of around

20-30 cases per 100,000 [47-49]. The increased levels of EBV related NPCs in these particular populations may be due to genetic predispositions, as it has been shown that individuals of Chinese decent, especially Cantonese males, are more likely to develop EBV-associated NPC regardless of where they live. Frequent genetic changes found in low-grade dysplastic lesions of high-risk individuals, in the absence of EBV infection, include deletions of regions on chromosomes 3p and 9p. This supports the argument that genetic changes may predispose to subsequent EBV infection [50, 51]. Cultural and dietary factors may also play an important role in the predisposition of these populations. This is supported by the fact that a decline in EBV-associated NPC incidence has been shown in second generation Chinese American migrants [13, 52, 53]. Dietary risk factors include preserved food preparation such as salted fish (especially during weaning), Qaddid (dry mutton preserved in oil) and the spice mixture Harissa [54].

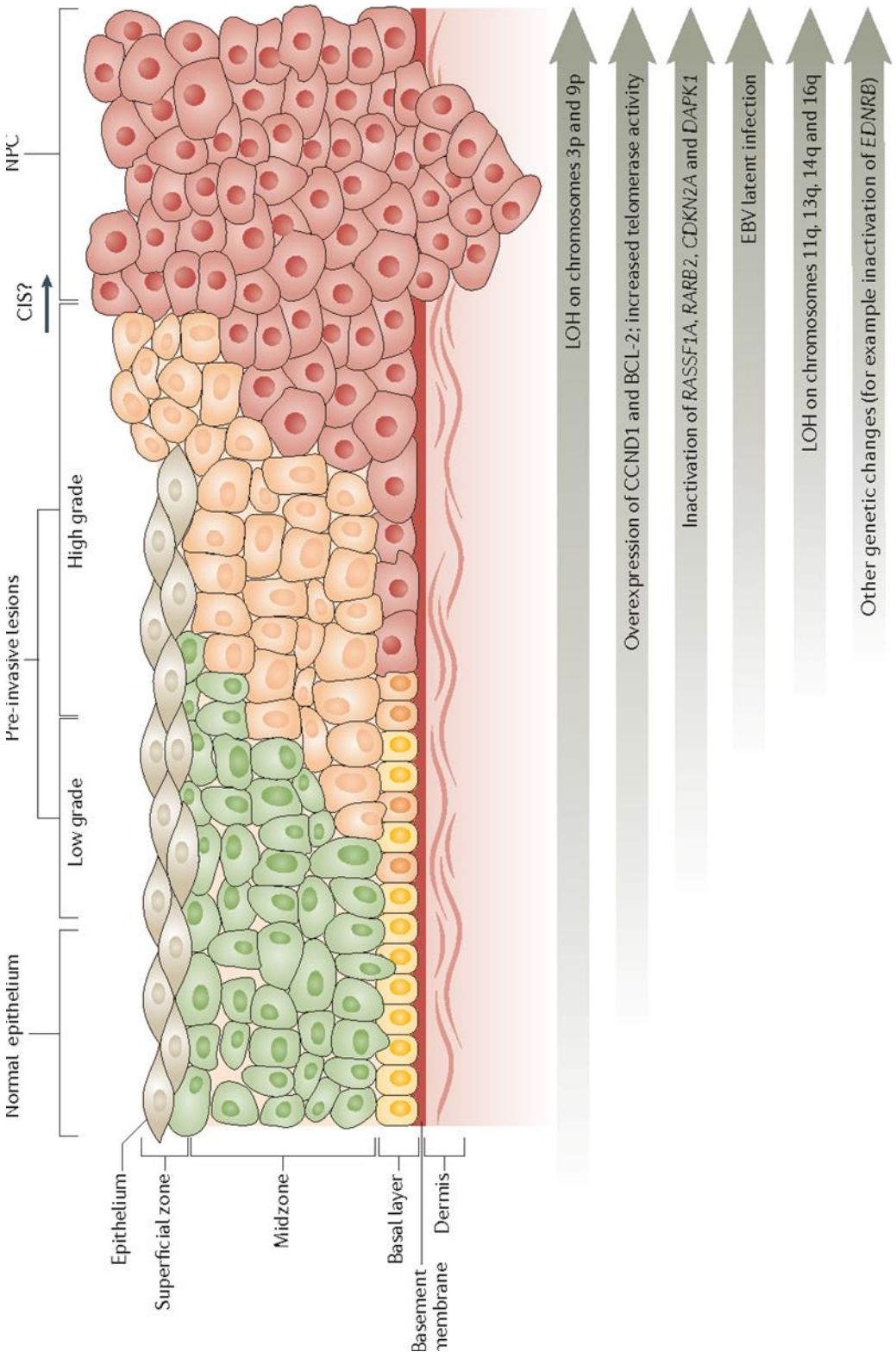


Figure 4: The pathogenesis of NPC. Different layers of the skin and cancerous states are indicated. Grey arrows show the carcinogenic changes thought to be associated with the above states [51].

It has been shown that EBV strains isolated from NPC tumours in endemic regions, such as M81, can infect primary epithelial cells and induce lytic cycle more efficiently than others strains (e.g. B95-8 or Akata) [55]. These strains often have mutations in HLA epitopes in LMP1 and proteins including gp110 and Zta [56, 57].

It is possible that NPC arises from persistently infected epithelial cells which have undergone a second genetic event. EBV causes the methylation and consequent silencing of multiple tumour suppressor genes in NPC [52]. Proliferation of epithelial cells may lead to other oncogenic transformations, as is the case in B lymphocytes. Another possibility is that EBV infection of epithelial cells occurs at a later stage of life in at-risk individuals whose nasopharynxes have been exposed to carcinogens or other preneoplastic genetic changes which render the affected epithelium susceptible to EBV infection. This argument is supported by the fact that EBV has been detected in the cells of the nasopharynx in high-grade pre-invasive lesions but not in low-grade disease [53, 54]. Thus strain variation, environmental, genetic and cultural factors likely all contribute to the development of NPC.

The immediate early lytic cycle protein Zta may enhances malignant transformation of epithelial cells by triggering the secretion of the angeogenic vascular endothelial growth factor and inflammatory, oncogenic cytokines (IL-6, IL-8, IL-10, IL-13) [55]. A prominent lymphocytic infiltration is characteristic of NPC tumours. The interaction between tumour cells and lymphocytes appears to be crucial for the continued propagation of the malignancy. However, the EBV genome is located in the epithelial tumour cells and cannot be detected in the infiltrating lymphocytes [44].

NPCs are broadly classified into two types: nonkeratinizing and keratinizing. Non-keratinizing NPC is by far the most common type, it can be poorly differentiated or undifferentiated and all tumour cells show evidence of clonal EBV infection. Non-keratinizing NPC is often accompanied by an ornate reactive T lymphoid response which may obscure the epithelial neoplasm leading to a lymphoepithelioma. However, the lymphoepithelioma does not show evidence of EBV infection [6]. Keratinizing NPC is defined by the presence of keratin in addition to the squamous appearance of the tumour cells. Initially, no association could be found between EBV and keratinizing NPC [56, 57], but recent studies have identified EBV in some cases of the disease [58]. Whether or not EBV plays a significant role in keratinizing NPC remains debatable.

While the exact point of EBV infection in NPC development is unclear, all cancerous cells in any EBV-associated NPC or high-grade dysplastic lesion carry monoclonal viral genomes, indicating that infection must occur before the clonal expansion of the malignancy, or that a single EBV infected clone has predominated [44, 59, 60]. Other EBV-associated neoplasms can arise in atypical cellular targets and include a subset of gastric carcinomas which share the histological features of undifferentiated NPCs [6, 15]. EBV infection may represent a rare event in these cell types.

1.3.5 Latency in Epithelial Cells

Based on the evidence to date, it has been suggested that true latency of EBV in epithelial cells only occurs in EBV-positive epithelial cell tumours [39]. Latency genes expressed in these cells appear intermediate between the latency 1 and

latency 2 programmes seen in B lymphocytes. EBNA1 and EBER are expressed in all EBV positive cases. Other EBNAs are absent. Some evidence points to growth-promoting effects of LMP2A on epithelial cells whose transcripts are amplified in most affected NPCs alongside LMP2B [25, 61]. The LMP2A protein can be detected in ~50% of cases. LMP1 protein can be identified in 35% of EBV associated tumours and is present in all pre-invasive lesions, suggesting an important role in early tumorigenesis that may not be required in established carcinomas [59, 61, 62]. It is still unclear how and when the virus infects epithelial cells, thus our understanding of the potential role of EBV in NPC development is limited. It is clear that the primary site of infection is the oral cavity and that oral secretions are responsible for the shedding and consequent transmission of the virus. However, it has not been possible to unequivocally identify EBV infection in oral epithelial cells of high-risk, healthy individuals; hence the occurrence of persistent infection in these cells is questionable [6, 15]. During *in vitro* experiments infection of B cells with EBV leads to the amplification of the EBV genome to>20 copies per cell, even if the infecting virus is a replication-defective Zta-knock out (see section 1.4 for further information about Zta). This amplification cannot be observed after infection of primary tonsillar epithelial cells (a likely site of EBN infection *in vivo*) in which EBNA1 expression is only detected in ~40% of EBER positive cells. The EBV genome is not maintained in these cells during culture. Conversely, the gastric carcinoma cell line AGS, which is derived from a malignant tumour formed of glandular structures in epithelial tissue, did result in maintenance of EBNA1 expression and EBV genome replication. However, as opposed to B cells, this replication was dependant on Zta. [39]

1.4 EBV Vaccines

There is currently no licensed prophylactic vaccine for EBV. Most efforts to develop an effective vaccine focus on the EBV glycoprotein gp350. Phase 2 clinical trials using monomeric gp350 showed a reduced susceptibility to infectious mononucleosis by 78% but did not reduce the rate of EBV infections. [63] Further phase 3 trials of the vaccine are needed to determine if the vaccine can indeed reduce the rate of infectious mononucleosis. Based on the fact that infectious mononucleosis increases the risk of developing Hodgkin lymphoma [64] and multiple sclerosis [65], a vaccine that reduces infectious mononucleosis might also reduce these associated diseases. However, due to the low rate of their occurrence it would require a considerably large study to determine this.

Due to the long period of time between initial EBV infection and the development of epithelial cancers like nasopharyngeal carcinoma or EBV positive gastric carcinoma, a clinical trial targeting these conditions would be difficult. A recent study showed that gp350 antibody levels correlate with a reduction in nasopharyngeal carcinoma development [66], so a vaccine that does not prevent EBV infection may still be effective against these malignancies. In addition, a study by deBruyn *et al.* showed that children with inflammatory bowel disease have low seropositivity to EBV [67], so it would be interesting to find out if an EBV vaccine has an effect on this condition.

Current efforts to develop an effective EBV vaccine include multimeric form of gp350, virus-like (nano-)particles and formulations including viral glycoproteins, lytic or latent proteins in addition to gp350. [68] Clinical trials on the effectiveness of such vaccines are still required.

The immediate-early lytic protein Zta, discussed in the following section, has been suggested as a vaccine target as it is produced prior to immune evasion genes. It is recognised more readily by CD8 T cells than early or late lytic proteins. [69] Zta is an important T cell target in infectious mononucleosis patients and patients with post-transplant lymphoproliferative disease. Destruction of cells expressing Zta reduces the chance of virion production. However, a recent study suggests Zta is not recognised in the early stages of EBV infection. [70]

Testing the efficacy of any EBV vaccines is complicated by the lack of a suitable non-primate animal model for EBV infection. In addition, difficulties in selecting a suitable adjuvant hinders progress in this field [14].

1.5 Zta

Zta (*alias* BZLF1, Z, ZEBRA or EB1) is a 245 amino acid, 35kDa, bZIP protein which is made up of an N-terminal transactivation domain, a basic DNA contact region, a dimerization region (zipper) and a c-terminal region [71],[72]. The activation of the Zp promoter, by AP1 proteins, Zta or SMAD [73], leads to the expression of BZLF1, which encodes Zta. It is the first viral gene to be expressed in the lytic cycle and coordinates the global changes in the viral transcriptome during the re-activation from latency as Zta activates further EBV lytic cycle genes [74]. It interacts with lytic promoters and origins of lytic replication, which are made up of a minimum of 32 sequence variants of a 7-mer sequence (ZREs), which resemble AP1 and C/EBPα binding elements. Most of the EBV genome, including the lytic promoters, is heavily methylated by the host cell during latency, creating a repressive chromatin environment [75]. Zta can recognize methylated

DNA and some CpG-containing Zta binding sites (CpG ZREs) are even dependent on methylation for Zta binding, which shows that Zta is able to interact with silenced DNA [76]. The loss of EBV genome methylation during lytic cycle leads to a reduction in binding of Zta to CpG ZREs, suggesting that Zta recognises these sites during the latency/lytic cycle transition, acting as a lytic cycle epigenetic switch [5, 77].

There are three different classes of ZREs. Class 1 is unmethylated as it lacks a CpG motif. Class 2 can be methylated and contains CpG and Zta binding is enhanced by methylation. It contains a variant cAMP response element binding protein (CREB) /AP1 site. Class 3 can also be methylated and contains CpG, but Zta binding dependant on methylation [78].

Previous chromatin precipitation investigations carried out in Group I EBV positive Akata BL cells have shown that Zta interacts with many sites on the viral genome including origins of lytic replication and the promoters of Zta, BRLF1, and BMRF1 [76, 78]. Based on experiments carried out in these cells, it is known that Zta also interacts with certain regions on the host cell genome [31, 79]. In addition, Zta interacts and co-associates with many host and viral proteins [80]. It forms a homodimer but cannot dimerise with other cellular bZIP proteins such as Fos, Jun, C/EBPa and the CREB [81]. However, Zta binds to different cellular transcription factors through its dimerization domain, including p65, p53 and the retinoic acid receptor [82-84]. Through poorly understood mechanisms, transfection of these proteins significantly inhibits the transcriptional activation function of Zta and thereby decreases activation of the lytic cycle [85, 86].

Zta stabilizes RNA polymerase II at Zta associated promoters [87]. It interacts with six of the seven viral replication proteins, except BALF2 [88]. Zta-DNA interaction may also be influenced by other factors such as co-associating host or viral proteins or chromatin structures.

An example of the Zta homodimer binding to an AP1 site is shown in Figure 5 below.

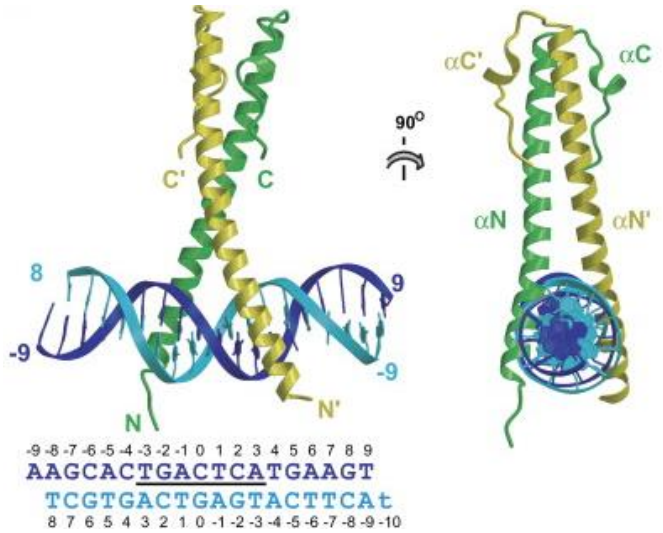


Figure 5: Model of Zta binding to DNA. One of the bZIP helixes (yellow) bends towards one end of the DNA while the other (green) is straight. The Ap-1 site has been underlined. Figure by Petosa *et al.*, 2006 [72].

The c-Fos, c-Jun, GCN4, and CREB protein family possess a conserved DNA-binding domain. Zta contains a similar basic domain and binds directly to the consensus AP1 site [89]. The Zp promoter contains an Ap1/CREB-like binding site which is the main auto-activation site for Zta. There are three CREB-specific co-activator transducer of regulated CREB activity (TORC or CRTC) proteins; of these, TORC2 is the best studied. During the lytic cycle, TORC2 is translocated

from the cytoplasm to the nucleus [90]. TORC1, 2 and 3 have been identified as co-activators of Zta on the Zp promoter [90]. One example is that, in reporter assays, co-expression of Zta and TORC2 enhances Zp transcription up to 108-fold in comparison to Zta expression alone [90]. This activation is complex. It is dependent both on ZREII (ZII) and ZREIII (ZIII) *cis* elements – ZREII (ZII) via CREB and ZREIII (ZIII) via Zta (Figure 5). siRNA knock-down of TORC2 reduces the levels of lytic proteins following induction, including Zta. Zta has been shown to form a complex with TORC proteins. In addition, both TORC2 and CREB have been shown to have some association with Zp. This interaction is significantly increased in the presence of Zta. TORC can associate with BCL3 which has been shown to recruit histone deacetylases that inhibit transcription [91] [92]. CREB has also been shown to be associated with Zp. Neither of these associations require Zta, but the effect of TORC2 is significantly higher when Zta is also present. The activation of Zp by TORC2 is increased by calcineurin, which is activated by calcium signalling following B cell receptor activation and leads to the dephosphorilation and consequent activation of TORCs. A model of the proposed mechanism of Zta interaction with TORC and CREB is shown in Figure 6 below [92].

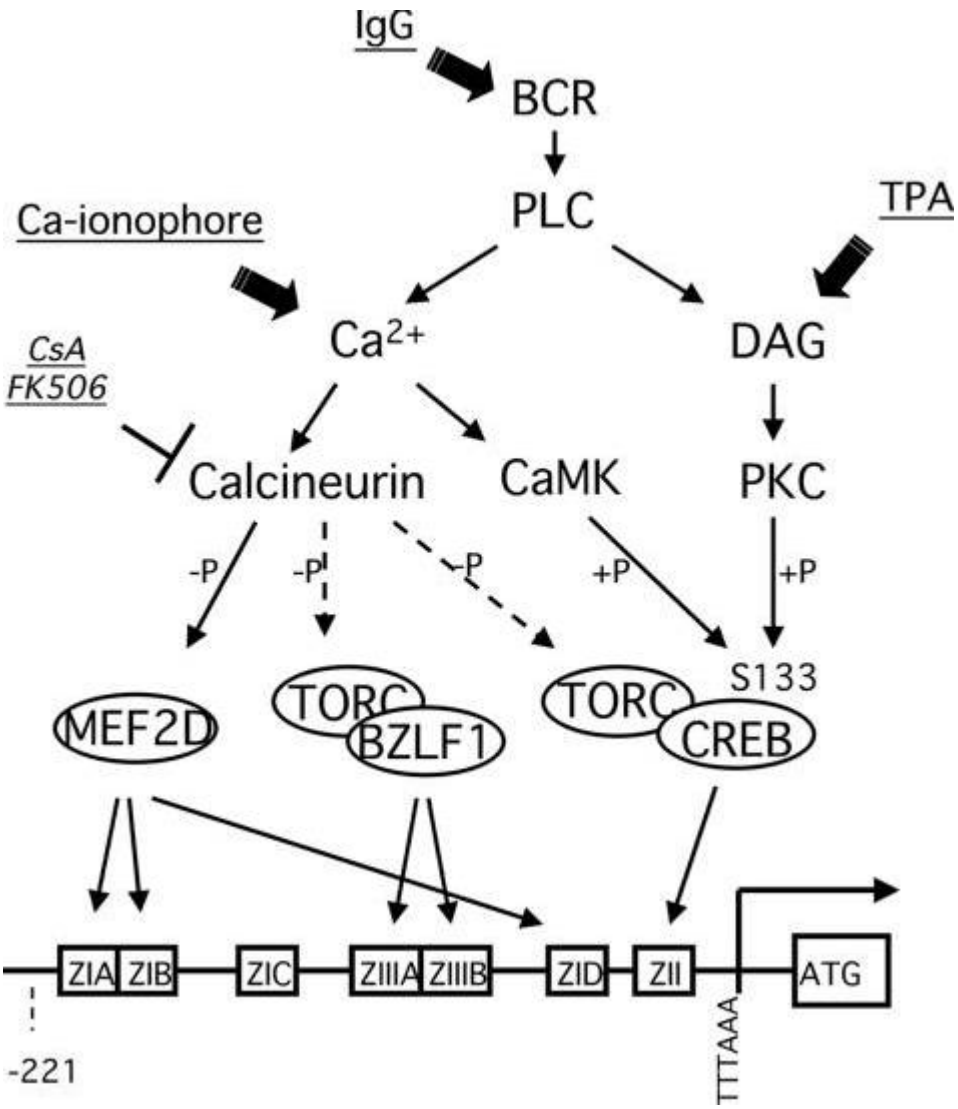


Figure 6: The proposed model for EBV Zp activation by Liang *et al.* 2002 [92]**.** TORC is activated by calcineurin. It is able to associate with CREB or Zta (BZLF1) and enhances Zp activity through binding to ZII and ZIII *cis* elements. PLC*:* phospholipase C, PKC*:* protein kinase C, DAG*:* diacylglycerol, CaM*:* calmodulin, BCR: B-cell receptor, MEF2D: myocyte enhancer factor 2D [90].

TORC is dephosphorylated in response to B cell receptor signalling which leads to its translocation into the nucleus where it increases the occupancy of pCREB

over certain binding sites [93]. It has been suggested that this allows for the specific activation of certain genes in response to signals which lead to both CREB phosphorylation and TORC dephosphorilation (Figure 7A). It is not clear if CBP is part of an alternative or cooperative pathways for target gene induction [94]. Nevertheless, TORC2 has been shown to interact cooperatively with CBP [95]. In addition, CBP has been shown to form a complex with Zta (and Rta [96]) and cooperates with it to significantly increase Zta-dependent transcription, including the lytic cycle, independent of its histone acetyltransferase activity. TORC2 has also been shown to interact with Zta and the TORC2/Zta complex is recruited to Zp. (Figure 7C) [97, 98]. In addition, Zta significantly increases CBP-mediated acetylation of nucleosomal histones [99]. CBP can also bind to cellular transcription factors including p65, p53 and the retinoic acid receptor. The association of these proteins with CBP correlates with transcription activation function [100, 101]. Interestingly p65, p53 and the retinoic acid receptor inhibit Zta transcriptional activity, and the over-expression of CBP relieves the antagonistic interactions between different classes of activators such as nuclear receptors and AP1 [100]. It has been suggested that competition for CBP may control the transcription output from a range of signal inputs. Figure 7B shows two potential mechanisms of this inhibition [98].

Zta has been shown to decreases the transactivation function of CREB in reporter assays. This effect can be partially reversed by the over-expression of CBP (Figure 7D) [97].

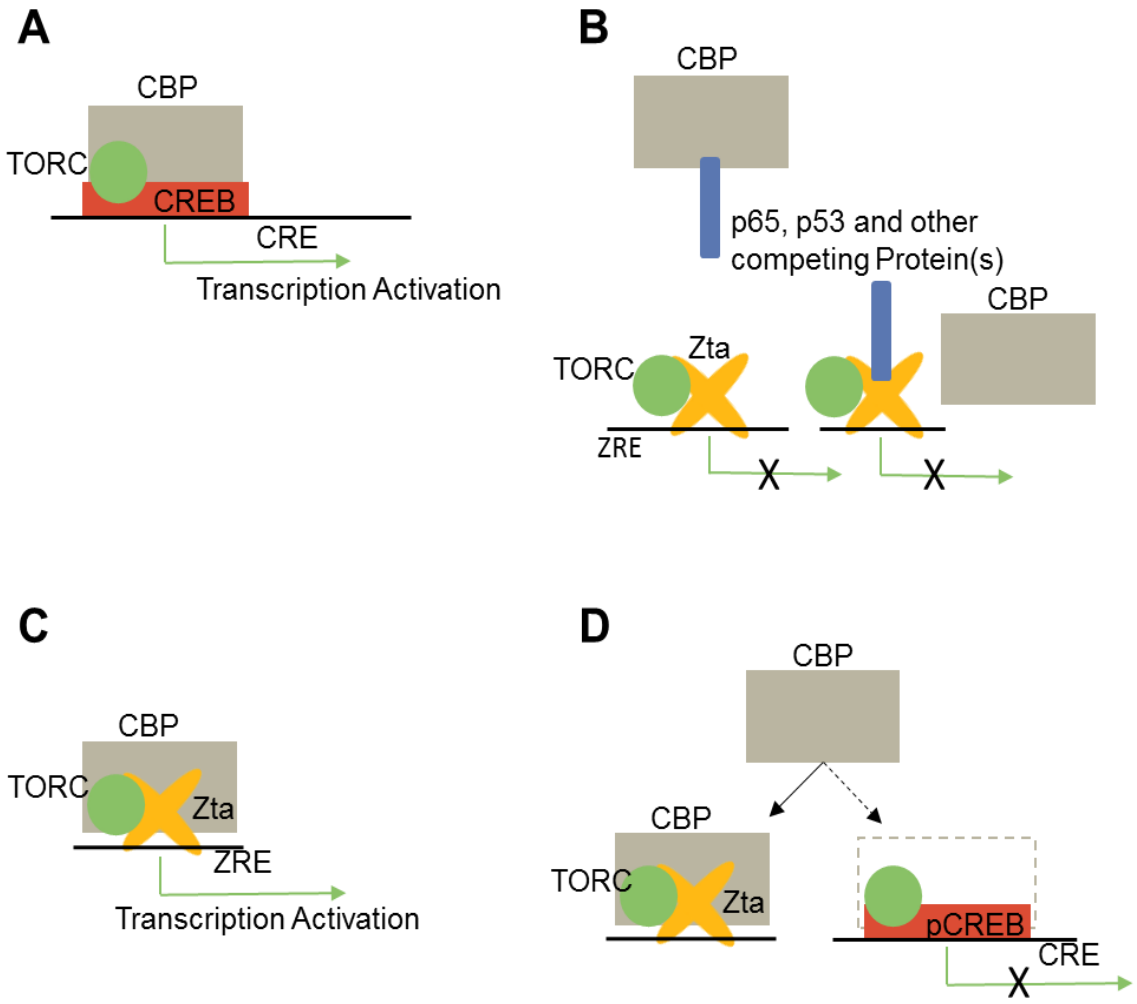


Figure 7: Model of cAMP response element binding protein (CREB) binding protein (CBP), Zta and phosphorylated CREB (pCREB) interaction. A: interaction of TORC, CBP and Zta increases Zta-dependant transcription of ZREs. B: Competition for CBP has been suggested to control transcriptional output of associated targets. Certain proteins may block binding of CBP to Zta by associating with CBP (left), alternatively competing proteins may associate with Zta thus hindering CBP-Zta interaction (right). C: TORC2 has also been shown to interact with Zta and the TORC2/Zta complex is recruited to Zp. D: Interaction of Zta with CREB decreases transactivation of CREB targets but has been suggested not to decrease the amount of pCREB-CBP complexes. Overexpression of CBP reverses this effect.

1.5 Project Context

My early work focussed on HONE1-EBV cells, originally described as a poorly differentiated nasopharyngeal carcinoma cell line, but now thought to be a HeLa derivative [102]. These cells are prone to lytic cycle reactivation and have been re-infected with recombinant Akata EBV strain, containing neomycin-resistance and GFP genes [103].

As mentioned previously, it has been shown that there may be certain epitheliotropic EBV subtypes which lead to spontaneous lytic replication of EBV in epithelial cells and which are associated specifically with EBV related epithelial cancer [104]. Sequence analysis of Akata and the epitheliotropic EBV subtype M81, showed complete sequence homology between Zta proteins in both EBV strains (Figure 8).

```
Akatagi|428161052|gb|AFY97859.      MMDPNSTSEDVKFTPDPYQVPFVQAFDQATRVYQDLGGPSQAPLPCVLWP 50
M81gi|557368273|gb|AGZ95183.1|      MMDPNSTSEDVKFTPDPYQVPFVQAFDQATRVYQDLGGPSQAPLPCVLWP 50
                                    **************************************************

Akatagi|428161052|gb|AFY97859.      VLPEPLPQGQLTAYHVSAAPTGSWFPAPQPAPENAYQAYAAPQLFPVSDI 100
M81gi|557368273|gb|AGZ95183.1|      VLPEPLPQGQLTAYHVSAAPTGSWFPAPQPAPENAYQAYAAPQLFPVSDI 100
                                    **************************************************

Akatagi|428161052|gb|AFY97859.      TQNQLTNQAGGEAPQPGDNSTVQPAAAVVLACPGANQEQQLADIGAPQPA 150
M81gi|557368273|gb|AGZ95183.1|      TQNQLTNQAGGEAPQPGDNSTVQPAAAVVLACPGANQEQQLADIGAPQPA 150
                                    **************************************************

Akatagi|428161052|gb|AFY97859.      PAAAPARRTRKPLQPESLEECDSELEIKRYKNRVASRKCRAKFKHLLQHY 200
M81gi|557368273|gb|AGZ95183.1|      PAAAPARRTRKPLQPESLEECDSELEIKRYKNRVASRKCRAKFKHLLQHY 200
                                    **************************************************

Akatagi|428161052|gb|AFY97859.      REVASAKSSENDRLRLLLKQMCPSLDVDSIIPRTPDVLHEDLLNF 245
M81gi|557368273|gb|AGZ95183.1|      REVASAKSSENDRLRLLLKQMCPSLDVDSIIPRTPDVLHEDLLNF 245
                                    *********************************************
```

Figure 8: Clustal 2.1 multiple sequence alignment of the Zta protein sequence of Akata and M81 EBV strains. No differences in the Zta protein could be detected. The top line is showing the Akata Zta protein sequence, the bottom line is showing the M81 Zta sequence.

1.6 Aims

My work aims to extend the knowledge surrounding the switch from latency to lytic cycle of the Epstein-Barr virus by:

- Identifying Zta binding sites in chromatin and making comparisons between epithelial and B cells
- Investigating histone modification and transcription factor binding in various cell types (ENCODE database) to identify common factors
- Exploring any relationship between cell factors and Zta function

Chapter 2: Materials and Methods

2.1 Materials

2.1.1 Cell Lines

Cell Line	Description	Reference	Application
Akata	EBV positive B lymphocyte, contains doxycycline inducible BZLF1 expression vector	[105]	Western blot, ChIP qPCR, RNA quantification, FACS, siRNA transfection
HEK 293T	EBV negative, derived from human embryonic kidney cells, contains mutant SV40 large T antigen	[106]	Luciferase assay, Western blot
HEK 293ZKO	derived from human embryonic kidney cells, infected with EBV lacking the BZLF1 gene	[31]	ChIP qPCR, Western blot, genome load, siRNA knock-down
HONE1-EBV	EBV positive adherent cell line	[107]	RNA quantification, western blot, FACS, genome load, ChIP qPCR, ChIP Sequencing

LCL3	EBV positive lymphoblastoid cell line, spontaneously lytic	[108]	Western Blot, FACS, ChIP qPCR, ChIP reChIP qPCR

2.1.2 Antibodies

Antibody	**Description**	**Supplier**	**Dilution**	**Application**
Actin	Rabbit polyclonal	Sigma (A2066)	1:5000	Western Blot
αIgG	Rabbit polyclonal	Dako (A0423)	1:800	Akata cell induction
Bcl-3 (150-3.5) FITC	Hamster monoclonal	Santa Cruz Biotechnology (sc-32741 FITC)	5µl to stain $1x10^6$ cells (see protocol for volume)	FACS
BCL3 (c-14)	Rabbit polyclonal	Santa Cruz Biotechnology (sc-185, discontinued)	1:200 or 10µl to stain $1x10^6$ cells	FACS and Western Blot
BCL3 (c-14) x	Rabbit polyclonal	Santa Cruz Biotechnology (sc-185 x, discontinued)	10µg/1ml IP or 1:2000	ChIP and Western Blot
BCL3 (new)	Rabbit Polyclonal	Abcam (ab27780)	1:500	Western Blot

ChIP negative control	Goat IgG	Santa Cruz Biotech (sc-2028)	10µg/IP (see protocol for volume)	ChIP control
IRDye 680CW	Fluorescent anti-mouse secondary	Li-Cor	1:5000	Western Blot
IRDye 800CW	Florescent anti-rabbit secondary	Li-Cor	1:5000	Western Blot
Mouse secondary Alexa Fluor 647	Donkey polyclonal	Abcam (ab150103)	5µl to stain $1x10^6$ cells (see protocol for volume)	FACS
normal Arm. hamster IgG-FITC	Hamster IgG	Santa Cruz Biotechnology (sc-2864)	20µl to stain $1x10^6$ cells (see protocol for volume)	FACS
negative control mouse	Mouse IgG	Sigma (I-8765)	1µl to stain $1x10^6$ cells (see protocol for volume)	FACS
negative control rabbit	Rabbit gamma globulin	Jackson Immuno Research (011-000-002)	0.2µl to stain $1x10^6$ cells (see protocol for volume)	FACS

Rabbit Secondary Alexa Fluor 647	Goat polyclonal	Abcam (ab150079)	5µl to stain $1x10^6$ cells (see protocol for volume)	FACS
TORC2	Rabbit monoclonal	Abcam (ab109081)	1:1000	Western Blot
TORC3	Rabbit monoclonal	Abcam (ab91654)	1:1000	Western Blot
VCA	Mouse monoclonal	Millipore (MAB8184)	1µl to stain $1x10^6$ cells (see protocol for volume)	FACS
Zta (BZ1)	Mouse monoclonal	Martin Rowe (Birmingham), described in [109]	1:500	Western Blot
Zta (EBV ZEBRA)	Goat polyclonal	Santa Cruz Biotechnology (sc-17503, discontinued)	1µl to stain $1x10^6$ cells (see protocol for volume)	FACS
Zta (SCZ)	Goat polyclonal	Santa Cruz Biotech (discontinued)	10µg/1ml lP (see protocol for volume)	ChIP

2.1.3 qPCR Primers

Primer	**5'–3' Sequence**	**Refe-rence**	**Appli-cation**
36B4	F: GGCCGAATRCTGTCTGTGGAGACCGATTACACC R:GGCCGGATCCGACTCTTCCTTGG CTTCAACCTTAG	[110]	RNA amount Normali-sation
BCL2A1	F: TCTTGAGCTGGCTCACCTTG R: AAACACAGCCTACGCACGAA	[111]	Zta and BCL3 ChIP qPCR
BCL2A1 3' flank	F: ACAGTGGTTACCTCTTGGGAGA R:CCTGTGTTGAAACTCATGTTGGTA	[111]	Zta and BCL3 ChIP qPCR
BCL2A1 5' flank	F: AGGAATTTGGCCTCCCAATCA R: TTTCTCCAGCGACCATGAGTT	[111]	Zta and BCL3 ChIP qPCR
beta globin	F: GGCAACCCTAAGGTGAAGGC R: GGTGAGCCAGGCCATCACTA	[112]	Genome load normalisation
BLLF3	F: TCCATCTGGGCACTTCTGACGCT R: CCGTCAGCAGCGTGTTCACAA	[18]	RNA quantify-cation
BNLF2a 5' Flank	F: CATGGCCCTGAACATGAGGT R: GAGGCCAGCCAAGTCTACAG	[18]	Zta and BCL3 ChIP qPCR
BNLF2a peak #2	F: GAGGCACAGCTGTTTTCGTG R: AGCCACACCTAACTCATGCC	[18]	Zta and BCL3 ChIP qPCR
BZLF1	F: AGCCAAGGCACCAGCCTCCT R: TGCATGAGCCACAGGCATTGCT	[18]	RNA quantify-cation

EBV polymerase	F:AGTCCTTCTTGGCTAGTCTGTTGAC R: CTTTGGCGCGGATCCTC	[112]	Genome load
EGFR #2	F: GCAGCAGCCTGTGAAACATT R: GCAACTGTTGGACTGGAGGA	n/a	Zta and BCL3 ChIP qPCR
EGFR 5' flank	F: TGGATGTGGGAACAGTGACG R: TCATGGCTTCCAGGATGCAG	n/a	Zta and BCL3 ChIP qPCR
FOSB	F:TTGCTAAGAAGATTTGCTCAGGTAG R: GCTTTACAATCTCCTTCTTACTGCT	[111]	Zta and BCL3 ChIP qPCR
GAPDH	F: CTCATGCCTTCTTGCCTCTT R: TTGATGGCAACAATATCCACTT	[113]	RNA amount Normali-sation
OriLyt Flank/BHLF1 OL1 flank	F: GCGCAACAGTGCCACCAACC R: CAGGACCTGGCGGTAGTGCAG	[76]	Zta and BCL3 ChIP qPCR
OriLyt/BHLF1 (OL5/OR4)	F: CAGCTGACCGATGCTCGCCA R: ATGGTGAGGCAGGCAAGGCG	[76]	Zta and BCL3 ChIP qPCR
SCIMP	F: CCCTCGTGCAATACTGTGAGA R: ACAACTCATTCGCTCTGGGC	[111]	Zta and BCL3 ChIP qPCR
	F: TTGCACAGCAAGTTCAAGCC	[111]	

SCIMP 3' flank	R: CTTTCTTGAAGGCAGATGGCAA		Zta and BCL3 ChIP qPCR
THRA #2	F: ACAGGGTCACACGTAGATGC	n/a	Zta and BCL3 ChIP qPCR
	R: GATGTCTGTGTGACTGGGCT		
THRA 3' Flank	F: AGCACTTCTCGCTATGCTCC	n/a	Zta and BCL3 ChIP qPCR
	R: TGTGTTCCCTGAGACCAAGC		

F: Forward, R: Reverse

2.1.4 TaqMan Primers

Primer Target	**Life Technologies TaqMan Assay ID**	**Reference**
ANO1	Hs00216121_m1	[111]
BCL2A1	Hs00187845_m1	[111]
BCL3	Hs00180403_m1	[114]
FOSB	Hs00171851_m1	[111]
FSCN1	Hs00602051_mH	[111]
GAPDH	Hs99999905_m1	[111]
GDF2	Hs00211913_m1	[111]
p50	Hs00765730_m1	n/a
p52	Hs01028890_g1	n/a
RASA3	Hs00183698_m1	[111]
SCIMP	Hs010294_m1	[111]
SLC6A7	Hs00204454_m1	n/a

2.1.5 Vectors

Name	**Backbone**	**Supplier/ Reference**	**Application**
BCL3 flag	pcDNA3	n/a	BCL3 expression
BHLF1	pCpGL-basic	[111]	Reporter of Zta function
BHLF1 mutant	pCpGL-basic	[111]	Negative control reporter vector
hisZta	pcDNA3	[115]	Zta expression
pcDNA3	pcDNA3	Invitrogen	Negative control expression vector

2.1.6 siRNAs

siRNA	**Supplier**
BCL3 (Attempt 2)	Dharmacon (M-003874-02-0005)
BCL3 1	Ambion (115617)
BCL3 2	Ambion (115619)
BCL3 3	Ambion (115618)
CBP/p300	Dharmacon (M-003477-02-0005)
Negative Control	Ambion (AM4611)
Negative Control (Attempt 2)	Dharmacon (D-001206-13-20)
Silencer Cy3 labelled negative control siRNA	Ambion (AM4621)
TORC1/CRTC1	Dharmacon (M-014026-02-0005)
TORC2/CRTC2	Dharmacon (M-018947-02-0005)

TORC3/CRTC3	Dharmacon (M-014210-01-0005)

2.1.7 Solutions and Buffers

Solution or Buffer	**Composition**	**Application**
Blocking buffer	1x PBS, 0.5% BSA (w/v)	ChIP for sequencing
Cell lysis buffer	85mM KCl, 0.5% (v/v) NP-40, 5mM PIPES (pH 8)	Chromatin preparation
FACS Wash	0.1% NaAzide, 5% FCS in PBS	FACS
High salt wash buffer	0.1% (v/v) SDS, 1% (v/v) Triton X-100, 2mM EDTA, 20mM Tris (pH 8), 500mM NaCl	ChIP
IP dilution buffer	0.01%(v/v) SDS, 1.1%(v/v) Triton X-100, 1.2mM EDTA, 16.7mM Tris (pH 8), 167mM NaCl	ChIP equilibration

LiCl wash buffer	250mM LiCl, 1%(v/v) NP-40, 1%(v/v) Na-deoxycholate, 1mM EDTA, 10mM Tris (pH 8)	ChIP
Low salt wash buffer	0.1% (v/v) SDS, 1% (v/v) Triton X-100, 2mM EDTA, 20mM Tris (pH 8), 150mM NaCl	ChIP
PBS-tween	138mM NaCl, 2.7mM KCL, 10mM Na2HPO4, 1.8mM KH2PO4 (pH 7.4), 0.1%(v/v) Tween-20	Western Blot
PMSF (100mM)	174mg Phenylmethylsulfonyl-fluoride in 10 ml of 100% ethanol	ChIP (at 1mM in each ChIP buffer)
SDS lysis buffer	1% (v/v) SDS, 10mM EDTA, 50mM Tris (pH 8)	Chromatin preparation
TE 1%SDS (Elution buffer)	10 mM Tris, 5 mM EDTA, 1%(v/v) SDS	ChIP

TE Buffer (Elution Buffer)	10 mM Tris, 5 mM EDTA	ChIP
Transfer buffer	25mM Tris-HCl (pH8.3),192mM glycine, 20% (v/v) methanol	Western Blot
Tris-Borate-EDTA (1xTBE)	89mM Tris (pH 7.6), 89mM boric acid, 2mM EDTA	DNA electrophoresis

2.1.8 Purchased Chemicals and Kits

Reagent or Kit	**Supplier**	**Application**
12% NuPAGE Bis-Tris SDS-PAGE Gel	Invitrogen	Protein electrophoresis
Agarose	Fisher Scientific	DNA electrophoresis
Agilent High Sensitivity DNA Kit	Agilent	ChIP sequencing library quality control
Agilent High Sensitivity DNA Reagents	Agilent	ChIP sequencing library quality control
Bioline 1kb hyperladder	Bioline	DNA electrophoresis
DharmaFECT reagent	Dharma-con	siRNA transfection
Dimethyl sulfoxide (DMSO)	Sigma	Cell Culture

DNA loading dye (5x)	Bioline	DNA electrophoresis
Doxycycline	Sigma	Induction of doxycycline promoter
Dried milk powder	Co-Op	Western Blot
Effectene Kit	HEK transfec-tion	Quiagen
Ethanol	Sigma	DNA purification
FIX and PERM Cell Permeabilization Kit	Invitrogen	FACS
Foetal calf serum (FCS)	GIBCO	Cell Culture
Formaldehyde	Sigma	Chromatin Preparation
G418	Fisher Scientific	Selection of EBV positive HONE1-EBV
Gel red	Biotium	DNA electrophoresis
Glycine	Sigma	Chromatin Preparation
GlutaMAX	GIBCO	NEON siRNA transfection
GoTaq Syber green master mix	Promega	qPCR
Laemmli Protein Sample Buffer (PSB, 2x)	Sigma	Protein electrophoresis
Luciferase reporter assay kit	Promega	Luciferase reporter assay
NEBNext ChIP Seq Library Prep Reagent Set for Illumina	New England Biolabs	ChIP Sequencing Library preparation

NEBNext Multiplex Oligos for Illumina (Index Primers Set 1)	New England Biolabs	ChIP Sequencing Library preparation
NEON 100µl kit	Invitrogen	NEON siRNA transfection
Morpholinepropanesulfonic Acid Buffer (20x MOPS)	NuPAGE	Protein electrophoresis
Odyssey blocking buffer	Li-Cor	Western blot
Passive lysis buffer (5x)	Promega	Luciferase Assay
PBS	GIBCO	Washing cells
Penicillin, Streptomycin, L-Glutamin (PSG)	GIBCO	Cell culture
Protein A Sepharose Beads	Sigma	ChIP
Protein G Sepharose Beads	Sigma	ChIP
Proteinase K	Sigma	DNA purification
Qiagen DNA Mini and Blood Mini kit	Qiagen	DNA purification
Qiagen MinElute Gel extraction kit	Qiagen	ChIP Sequencing Library preparation
QIAquick PCR Purification Kit	Qiagen	DNA clean-up
Qiagen RNeasy Plus Kit	Qiagen	RNA purification
Qubit dsDNA HS Assay Kit	Life Techno-logies	ChIP
RPMI 1640 media	GIBCO	Cell Culture

ROCHE complete mini Tablets	ROCHE	ChIP
Suberoylanilide Hydroxamic Acid (SAHA)	Sigma	Lytic cycle induction
Salmon sperm DNA	Invitrogen	ChIP
SeeBlue protein marker	Invitrogen	Protein electrophoresis
siRNA buffer (5x)	Dharma-con	siRNA transfection
SYBR gold	Life Techno-logies	DNA electrophoresis
Transcription first strand cDNA synthesis kit,	Roche	cDNA synthesis
Trypsin-EDTA (0.25%), phenol red	GIBCO	Detaching tightly adherent cells

2.2 Methods

2.2.1 Cell Culture and Induction of EBV lytic Replication

Every cell line used in this thesis was maintained in RPMI medium supplemented with 10% (vol/vol) fetal bovine serum (FBS), 100U of penicillin/ml, 100g of streptomycin/ml, and 2mM L-glutamine (Invitrogen) at 37°C with 5% CO_2.

HONE1 cells were kindly provided by G. Tsao. Transfection of HONE1 cells with EBV is not stable. The EBV plasmid contains a G418 resistance gene. EBV

positive HONE1-EBV cells were selected with 600µg/ml G418. For EBV lytic induction, HONE1-EBV cells were induced with 10µM suberoylanilide hydroxamic acid (SAHA) for 48hrs at 70% confluency. For EBV lytic cycle induction, Akata group I EBV-positive BL cells were seeded in log phase growth. After 24hrs the cells were concentrated and treated with 0.125% rabbit anti-human IgG (DAKO) for a further 24hrs. Alternatively, lytic cycle in Akata cells which contain the bidirectional doxycycline inducible promoter B-Tet, which drives expression of BZLF1 and green fluorescent protein (GFP), was induced for 24hrs with 500ng/ml doxycycline, a kind gift from Rajaei Almohammed and Ijiel Barak Naranjo Perez Fernandez. Lytic cycle in HEK 293ZKO cells was induced by transfection of a hisZta expression plasmid.

Akata and LCL3 cells were grown in suspension. HONE1-EBV and HEK cells are adherent cells. When harvesting, HEK 293T and HEK 293ZKO cell were detached from the flask with PBS. HONE1-EBV cells were detached using trypsin, or by scraping for chromatin preparation.

2.2.2 DNA extraction

Cells were re-suspended in PBS (at 37°C) and DNA was extracted using the Qiagen DNA Mini and Blood Mini kit (50) and accompanying protocol. The waste collection tube was replaced after each centrifugation. The DNA was eluted in 50µl AE buffer and stored at -20°C or used for downstream experiments immediately.

2.2.3 Quantitative Polymerase Chain Reaction (qPCR)

qPCR analysis was undertaken using the GoTaq qPCR master mix from Promega on a StepOnePlus real-time PCR machine (Applied Biosystems). Each reaction contained 12.5µl GoTaq (Cybr green), 0.5µl primer mix (5µM each forward and revers), 9.5µl PCR grade water and 2.5µl sample. Cycle conditions were: 10min at 95°C, followed by 40 cycles of 15sec at 95°C and 1min at 60°C. Data was collected and analysed with the help of the StepOne software version 2.3 (Applied Biosystems). Resulting values are relative to a standard curve made by serial dilution, or a reference sample for some RNA quantifications. For ChIP qPCR, 4% of input, or serial dilutions of sample DNA were used to generate a standard curve for absolute quantification.

2.2.4 Western Blot

Cells were harvested and cells (or chromatin) were boiled in protein sample buffer (Sigma) at 95°C for 5min. Transfected samples were sonicated for six rounds of 10sec pulses at 21% amplitude using a Vibra-Cell Sonicator (SONICS). Samples were loaded at 8-10µl per well and the protein was resolved alongside 5µl of SeeBlue marker ladder (Invitrogen) on a NuPAGE 12% Bis-Tris polyacrylamide gel (Life Technologies) and run in 1x MOPS buffer (Invitrogen) at 200V for 50min. The proteins were transferred onto nitrocellulose membranes (Santa Cruz Biotechnology), in transfer buffer, by placing the gel next to the membrane inside of filter papers and sponges in a cooled Bio-Rad Transfer tank at 75V for 90min. The nitrocellulose membrane was blocked in 5% milk for 30min, then incubated overnight with the relevant antibodies at 4°C in 5% milk. The membrane was washed three times for 5min in PBS-tween at room temperature to remove the

excess primary antibody. The relevant IRDye secondary antibody was added to Odyssey blocking buffer diluted 1:1 in PBS. This was added to the membrane for 2hrs at room temperature. Following three 5min washes at room temperature in PBS-tween, the western blot was visualised by chemiluminescent on the Li-Cor on the 700nm and/or 800nm channel with an exposure time of 2min. All incubation steps during the processing of the membrane were undertaken on a rocker.

2.2.5 Flow Cytometry and Fluorescence-Activated Cell Sorting (FACS)

Cells were prepared by fixation, permeabilization and staining as indicated using the manufacturer's guidelines of the Fix and Perm Cell Fixation and Cell Permeabilization Kit by Life Technologies. Cell were sorted using the BD FACSCanto analyser (Beckton Dickinson) or the BD Accuri C6. Data was analysed and graphs generated with FlowingSoftware or FCS Express 5 Flow (DeNovo Sofware) respectively.

2.2.6 Chromatin Immunoprecipitation (ChIP)

Chromatin was fixed by application of 1% formaldehyde for 15min, or 10min for the chromatin sonication protocol II, at room temperature.0.125M glycine was used to halt the cross-linking. Cells were pelleted at $1x10^7$ cells per ChIP and washed in PBS before being resuspended in 600µl cell lysis buffer. Cells were incubated for 10min at room temperature and pelleted at 3500rpm for 10min at 4°C. The pallet was resuspended in 200µl ChIP SDS lysis buffer.

For the original chromatin sonication protocol, cell lysates were pooled to a volume of 2ml and sonicated on ice for ten rounds, with 10sec pulses and a 30% amplitude output on a Vibra-Cell 750 Sonicator (SONICS). For the improved chromatin sonication protocol II, cell lysates were sonicated in a volume of 100µl on ice for ten rounds, with 10sec pulses and a 21% amplitude output on a Vibra-Cell 750 Sonicator (SONICS). Samples were vortexed after every minute. Sonication efficiency was checked by running the sample on a 1% agarose gel with 1x TBE buffer.

100µl of mixed protein A and protein G beads of 50% (w/v) bead slurry was used per immune-precipitation. Beads were prepared by washing and re-suspending in 55µl ChIP buffer for ChIP qPCR, or blocking buffer for ChIP sequencing. Beads were blocked for 3min at 4°C with 55µl 10mg/ml salmon sperm for ChIP qPCR or blocking buffer for sequencing. Beads were then washed and re-suspended in 55µl IP dilution buffer.

Chromatin samples were diluted 1:10 in IP dilution buffer and pre-cleared with 45µl of beads at 4°C for 30min. Samples were centrifuged and the supernatant was transferred to a new tube. 40µl was removed from the sample and frozen at -20°C to act as 4% input control.

For each IP, 20µg antibody was used with chromatin from $5x10^6$ cells. Samples were incubated at 4°C for 1hr while rotating. Immune complexes were then collected over night with 55µl 50% pre-blocked protein A/G-sepharose beads at 4°C.

Beads were spun down at 3000rpm for 5min at 4°C and washed with low salt, high salt, LiCl, and TE (2x) wash buffers for 15min at 4°C while rotating. The

antibody-bound complexes were eluted from the beads with ChIP elution buffer at 65°C for 15min. Input samples were thawed and included in further processing. To reverse crosslink the chromatin from the proteins, supernatant was incubated at 65°C over night. Samples were diluted 1:2 in TE buffer and 2.5µl proteinase K. Samples were incubated at 55°C for 3hrs. DNA was purified using a PCR purification kit (Qiagen) and eluted in 55µl. ChIP sequencing was performed using a 7x immunoprecipitation protocol, finally eluted into 45µl to achieve a sufficient DNA concentration of >10nM.

2.2.7 ChIP reChIP

The ChIP reCHIP protocol starts off like the ChIP protocol in section 2.2.6. After eluting the antibody-protein-DNA from the protein A beads at 65^0C for 15min with TE/1%SDS, retaining the supernatant and adding TE/1%SDS to the defrosted input control, the ChIP protocol was repeated for the second round of ChIPs by following the steps from 1:5 sample dilution, with IP dilution buffer containing protease inhibitor and pre-clearing, to the end of the CHIP protocol.

2.2.8 ChIP-Sequencing Library preparation

The ready to sequence libraries were prepared according to the NEBNext ChIP-Seq Library Prep Reagent Set for Illumina kit instructions (New England Bioscience), switching the PCR and the size selection steps. As is described in the workflow in Figure 9, the sonicated DNA fragments (1) were end repaired and the 5’ ends were phosphorylated and dA-tailed (2) to allow for adaptor ligation (3). U excision of the adaptor was performed to linearize the barcode binding region and to allow the PCR reaction to take place (4). Barcodes (Index 4, 6 and

12) and primers were ligated and PCR was performed to amplify the DNA (5). Barcodes make it possible to sequence several different samples during the same run. Size selection was performed on a 2% agarose gel with Gel Red (Biotium) or SYBR Gold dye (Life Technologies) to exclude the adaptor from the sequencing process. DNA fragments in the 175-225bp size range were excised and purified using the Qiagen Gel Extraction kit (6).

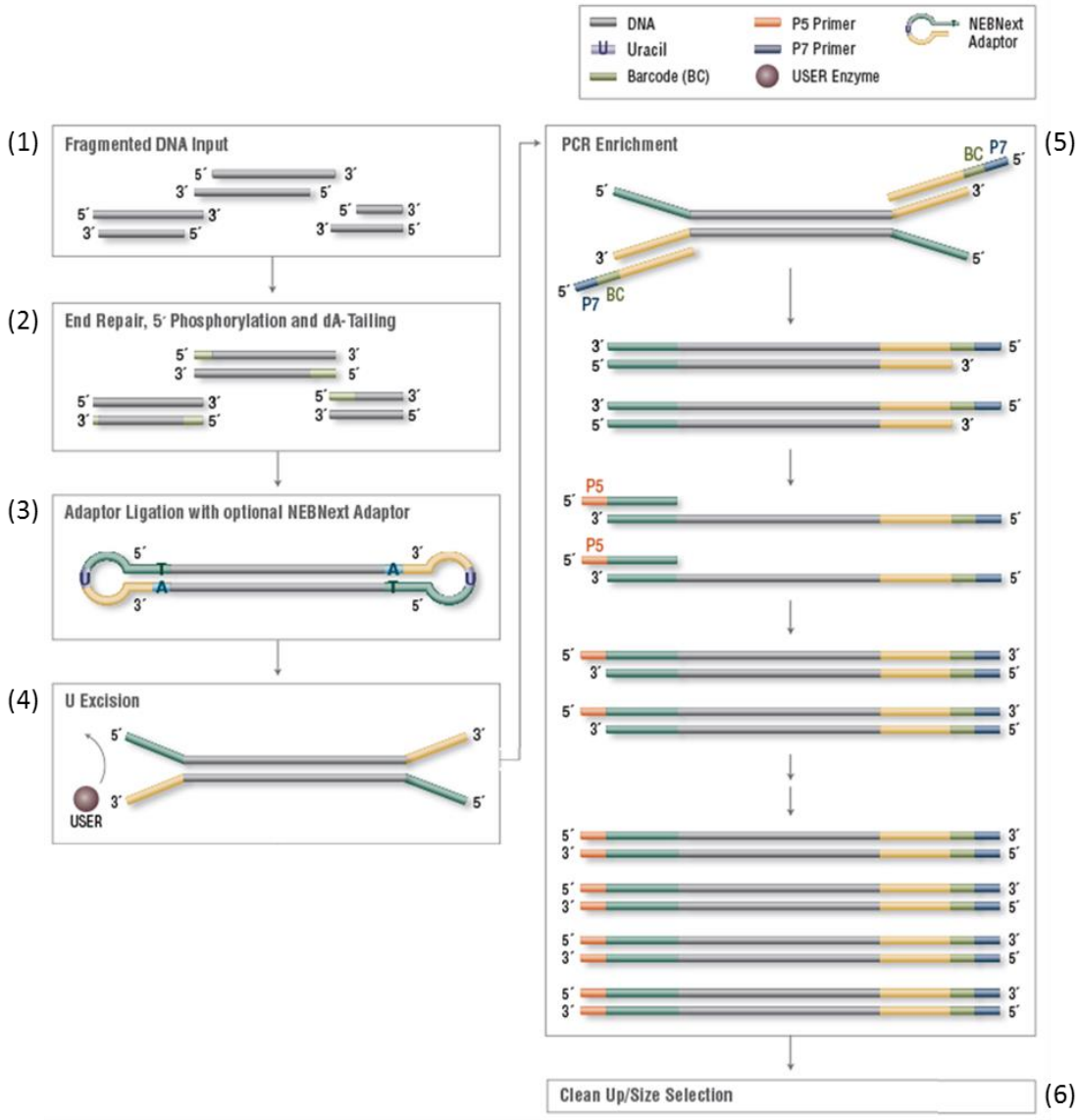


Figure 9: DNA Library Preparation Workflow for Illumina. Figure modified from NEBNext® DNA Library Prep Master Mix Set for Illumina® product information (https://www.neb.com/products/e6040-nebnext-dna-library-prep-master-mix-set-for-illumina)

2.2.9 ChIP Library Quality Control

Quality control was performed on the Agilent 2100 Bioanalyzer. Samples were prepared for this purpose following the instructions of the Agilent High Sensitivity DNA Bioanalyzer reagents kit. Samples were sequenced by the Eurofins Genomics group using Illumina HighSeq 2000 technology.

2.2.10 ChIP sequencing Data Analysis

Sequence reads were mapped to the human (GRCh37.p13) and EBV (NC_007605.1 GI:82503188) reference genomes by the Eurofins Genomics group. The bamFingerprint graph was generated with the help of the deepTools Galaxy platform (http://deeptools.ie-freiburg.mpg.de/) [116] to the parameters stated in Table 2.

Input Parameter	**Value**
Length of average fragment size	200bp
Region of genome to limit the operation to	none
Bin size	30bp
Number of Samples	1000000
Minimum mapping quality	1

Table 2: bamFingerprint Parameters

Binding of Zta to the human genome was visualised in the UCSC genome browser (http://genome.ucsc.edu/) [117]. For this purpose, the *.bam* files were

normalize to fragments per kilobase per million reads and converted to *.bigwig* format with the help of the bamCoverage tool in deepTools Galaxy. Parameters for this analysis were: 200bp fragment size and 30bp bin size.

MACS peaks were determined using the MACS-1.4.2 tool in deepTools Galaxy using the parameters outlined in Table 3.

Input Parameter	**Value**
Control data	Input
Tag size	30bp
Bandwidth	200bp
p-value	10^{-5} - 10^{-9}
Do not built shifting model, arbitrary shift size	100bp
Format	*.wig*
MFOLD range	10,30
Dynamic lambda	1000 - 10000

Table 3: MACS peak Parameters

MACS peaks were intersected using this “Operate on Genomic Intervals” option in deepTools Galaxy. Overlapping pieces of intervals were selected as return format (see Figure 10).

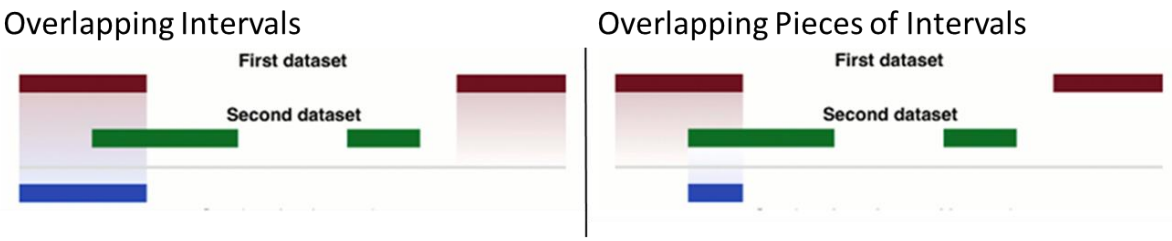


Figure 10: Return file options for the intersection of two MACS peak datasets. Return files are shown in blue. Figure modified from http://deeptools.ie-freiburg.mpg.de/.

Motif analysis was performed with the help of the MEME-ChIP suite using default parameters (http://meme.nbcr.net/meme/tools/meme-chip) [118].

Gene list enrichment analysis was performed using the EnrichR tool (http://amp.pharm.mssm.edu/Enrichr/) [119].

Associated genes were determined and analysed using the Genomic Regions Enrichment of Annotations tool, GREAT (http://bejerano.stanford.edu/great/public/html/) [120]. For this purpose, *.bed* files were created by the USCS genome browser from the MACS peak data sets and sent directly from the genome browser to GREAT.

Enrichment of Zta binding sites on the EBV genome was visualised using the Integrative Genomics Viewer (https://www.broadinstitute.org/igv/) [121] and the reference genomes for EBV (NC_007605.1 GI:82503188).

BCL3, EZH2, SUZ12 and H3K27me3 ChIP sequencing data was obtained, through the UCSC genome browser, from the ENCODE Transcription Factor Binding Tracks database. Track and file names are listed in Table 4.

ChIP site	**Track**	**File Name**
BCL3	GM12878 TFBS Uniform Peaks of BCL3 from ENCODE/HudsonAlpha/Analysis	wgEncodeAwgTfbsHaibGm12878BCL3V0416101UniPk.narrowPeak.gz
EZH2	GM12878 TFBS Uniform Peaks of EZH2_(39875) from ENCODE/Broad/Analysis	wgEncodeAwgTfbsBroadGm12878Ezh239875UniPk.narrowPeak.gz
SUZ12	H1-hESC TFBS Uniform Peaks of SUZ12 from ENCODE/USC/Analysis	wgEncodeAwgTfbsSydhH1hescSuz12UcdUniPk.narrowPeak.gz
H3K27me3	GM12878 H3K27me3 Histone Mods by ChIP-seq Peaks from ENCODE/Broad	wgEncodeBroadHistoneGm12878H3k27me3StdPkV2.broadPeak.gz

Table 4: BCL3, EZH2, SUZ12 and H3K27me3 ChIP sequencing data track and file names as stated in the UCSC genome browser.

2.2.11 RNA Extraction

Surfaces and equipment was cleaned with RNAse AWAY spray (Thermo Scientific) prior to RNA extraction. Cells were pelleted to $5x10^6$ and the pellet was thoroughly loosened by vortexing. RNA was extracted from cells using the RNeasy Plus Mini Kit from Qiagen following the kit'sprotocol. All centrifugation

steps were performed at 20–25ºC in a standard microcentrifuge, ensureing that the centrifuge does not cool below 20ºC. Prior to transfer of the homogenized lysate to a gDNA Eliminator spin column, the lysate was heated to 37ºC. To elute the RNA, 32µl RNase-free water was added to the column and incubated at room temperature for 10min before centrifugation. The RNA concentration and purity were measured on the Nanodrop at A260 (see 2.2.12). RNA was stored at -80ºC or used for downstream applications immediately.

2.2.12 RNA Quantification

RNA was extracted from cells using the Quiagen RNeasy Plus Kit. The concentration of the purified RNA was measured via spectrophotometry on the Nano Drop 2000 (Thermo Scientific) and transcribed to cDNA using the Transcription First Strand cDNA Synthesis Kit (Roche, see 2.2.12).

2.2.13 cDNA Synthesis

Surfaces and equipment was cleaned with RNAse AWAY spray (Thermo Scientific) prior to cDNA synthesis. RNA was reverse transcribed using the Transcription first strand cDNA synthesis kit (Roche) with its accompanying protocol using random hexamer primer. cDNA was stored at -20ºC or used for qPCR immediately.

2.2.14 NEON siRNA transfection

Cells were seeded in log phase and transfected 24hrs later in antibiotic free media containing 1% GlutaMax (GIBCO) using the manufacturers guidelines of

the NEON transfection system 100µl kit with buffer R. Pulse conditions were set as follows: Pulse voltage: 1300V, Pulse width: 30ms, Pulse number: 1. siRNA concentration varied and are stated with the corresponding result.

2.2.15 DharmaFECT siRNA transfection

siRNA transfection was performed as described in the manufacturer's DharmaFECT Transfection Reagent siRNA Transfection Protocol for HEK 293 cells with 1x siRNA buffer and a final siRNA concentration of 25nM. TORC1, TORC2 and TORC3 were transfected together at a concentration of 25nM each. Cells were transfected with hisZta or pcDNA3 24hrs after siRNA transfection and harvested a further 24-48hrs later. Samples harvested after 24hrs were also treated with 1µM ionomycin and 20ng/ml Phorbol 12-myristate 13-acetate (PMA) at the time of hisZta/pcDNA3 transfection.

2.2.16 Vector Transfection

Expression and reporter vectors were a kind gift from Kay Osborn. Cells were seeded into log phase growth 24hrs prior to transfection. Transfection was performed using the non-liposomal Effectene transfection kit (Qiagen) following the manufacturer's guidelines, except the amount of effectene was altered to 2.5µl per µg DNA. The amount of vector DNA added per experiment is shown with the result. HEK 293T cells were harvested 48hrs after transfection to perform the luciferase assay. HEK 293ZKO cells were harvested 24-72hrs after transfection to measure protein levels and EBV genome load.

2.2.17 Luciferase Assay

HEK 293T cells were transfected with hisZta, BCL3 or pcDNA3 control expression plasmids together with BHLF1 or BHLF1 mutant luciferase reporter vectors as described in section 2.2.11. Cells were harvested in PBS after 48hrs. Some of the cells were used for a corresponding western blot. The other cells were processed using the Luciferase reporter assay kit by Promega. Cells were lysed in 1x passive cell lysis buffer (Promega), spun down and the supernatant was moved to a new tube. 10µl of each sample was pipetted into a white 96-well plate and 50µl of LAR (Promega) was added to the samples and bioluminescence was measure for 10sec with a 2sec delay by the Promega GloMax multi-detection system. The background fluorescence was subtracted using measurements in wells containing only lysis buffer.

Results were normalised with the help of a bicinchoninic acid (BCA) assay to account for differences in overall protein levels across samples. For this purpose, a standard curve of known Bovine Serum Albumin (Pierce) concentration was created. The standard curve and sample lysates were added to a white 96-well plate followed by 200µl of working reagent (Pierce). After 30min incubation at 37°C, the absorbance of each sample was read at 560nm on the Promega GloMax multi-detection system. Cell-free background readings were subtracted.

Chapter 3: Zta binds to the human and viral genome in epithelial cells

3.1 Introduction

In order to study EBV in epithelial cells, HONE1-EBV cells were used, which, at the time, were thought to derive from a poorly differentiated NPC tumour biopsy. However, shortly after key experiments were performed, it was shown that HONE1-EBV contains HeLa genomic material [102] and a variant of STR loci from HeLa [122]. This suggests that HONE1-EBV may consist of a hybrid HeLa as has previously been suggested for other NPC cell lines [123]. More research is still needed to identify the origin of the HONE1-EBV cell line. That being said, HeLa cells are an epithelial cell line and HONE1-EBV is infected with EBV, so results obtained from this study are relevant in extending the knowledge surrounding EBV lytic cycle in epithelial cells.

Based on unpublished data from the Tsao laboratory (University of Hong Kong), who kindly provided our cells, EBV lytic cycle in HONE1-EBV cells was induced with 10μM suberoylanilide hydroxamic acid (SAHA, Vorinostat) for 48 hours. The exact mechanism by which SAHA induces lytic cycle is not known. It is a potent histone deacetylase inhibitor which facilitates the transcription of genes that result in apoptosis, differentiation and growth arrest [124]. SAHA is used in the treatment of cutaneous T cell lymphoma and is currently in Phase II clinical trials for the treatment of nasopharyngeal carcinoma [125, 126]. In addition it is widely used for research purposes to study gene regulation, transcription regulation and cell signalling.

Zta coordinates the change during the switch from latency to lytic cycle of EBV, as such, it is one of the earliest indicators of lytic replication [127]. To verify our

method of lytic cycle induction, we looked for Zta expression in cells treated with SAHA. We also examined the expression of a second early lytic cycle gene, BLLF3, which encodes a UTPase [128].

Once it was confirmed that the viral lytic cycle could be induced successfully, chromatin was prepared and sonicated to create fragments of 200-600bp for chromatin immunoprecipitation to ask whether Zta binds at the viral origin of lytic replication promoter (OriLyt, BHLF1) and the BCL2A1 promoter of the human genome, which encodes an anti-apoptotic BCL2 family member [129]. These loci are known Zta binding sites in Akata B cells [19].

Based on the obtained results, a ChIP sequencing library was prepared. Following optimisation of the protocol, the samples were shown to be of suitable quality and sent to an external laboratory for sequencing and read alignment to the EBV and human genomes.

An example of Zta binding to the viral genome with a clear peak around OriLyt in duplicate ChIP experiments is presented. Downstream data analysis focussed mainly on Zta binding to the human genome and is outlined in more detail in chapter 4.

3.2 EBV Lytic Cycle Induction in HONE1-EBV cells using SAHA

In order to find an effective way of inducing EBV lytic cycle in HONE1-EBV cells, induction condition received from the Tsao laboratory, who kindly provided the cell line, were tested. This involved investigation of the amounts of Zta protein, through western blotting and FACS. Zta RNA levels were examined with the help

of reverse transcribed complementary DNA (cDNA) quantitative polymerase chain reaction (qPCR). The EBV genome copy number per cell was investigated via qPCR.

Zta (BZLF1) and BLLF3 RNA quantification was carried out, which showed an eight-fold increase in Zta RNA levels and nearly 17-fold increase in BLLF3 transcription relative to the amount of 36B4, which was used as a cellular reference gene (Figure 11).

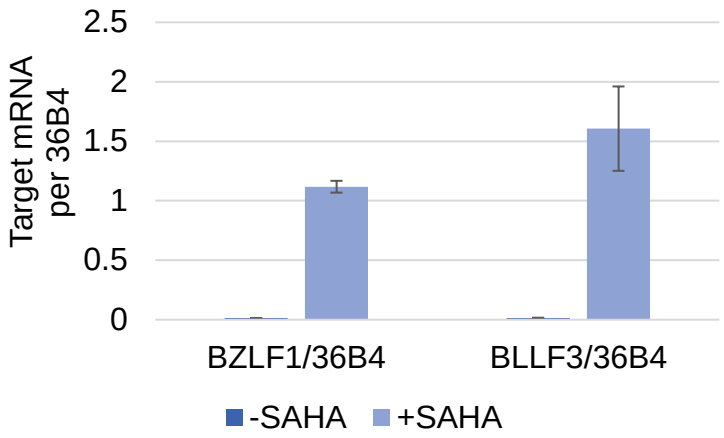


Figure 11: qPCR result showing the amounts of Zta (BZLF1) and BLLF3 RNA relative to host cell 36B4 levels. EBV lytic cycle was induced (+SAHA) or unindicted (-SAHA) with 10μM SAHA for 48 hours. Values are represented as mean ± SD. The amount of BZLF1 and BLLF3 RNA is significantly higher in samples treated with SAHA in comparison to untreated samples ($p<0.05$).

The increased Zta RNA amount translates into an increase of Zta protein, Zta expression in induced cells was determined via western blotting (Figure 12). This

showed Zta in induced cells. In contrast, no Zta can be detected in the uninduced control.

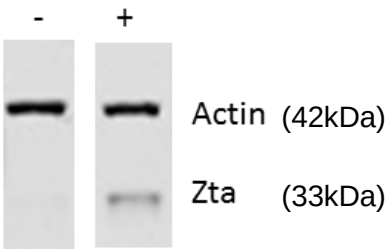
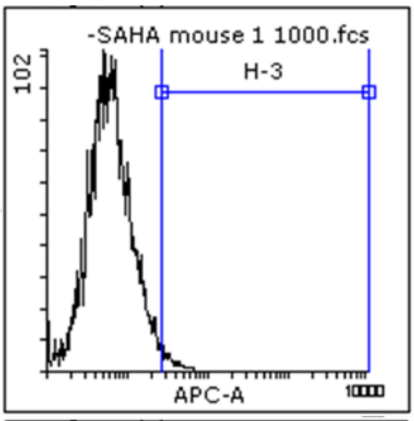


Figure 12: Western blot showing the expression of Zta in HONE1-EBV cells. Lytic Cycle was induced (+) or uninduced (-) with 10μM SAHA for 48 hours. An antibody specific for actin was used as a loading control.

In addition to western blot analysis of Zta levels, Fluorescence-activated cell sorting (FACS) was also carried out to detect the percentage of cells expressing the Zta protein. Results are summarised in histograms shown in Figure 13. No clear difference between cells induced with SAHA and uninduced cells could be detected during this analysis. Cells that are Zta positive lie within the gated region of the histogram, which accounts for 6.8% of uninduced cells (-SAHA) and 7.6% of induced cells (+SAHA).

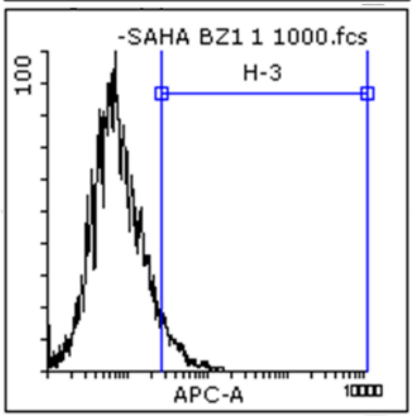
A



B

Experiment	Zta positive (%)
Mouse control antibody, not induced	1.98
Zta antibody, not induced	6.84
Zta antibody, induced	7.64

C



D

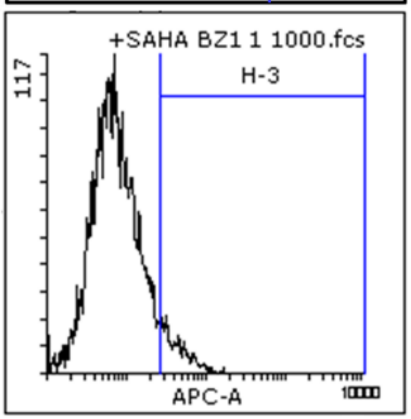


Figure 13: Histograms showing FACS analysis results of Zta levels. EBV lytic cycle in HONE1-EBV cells was induced with 10μM SAHA (+SAHA, D) or uninduced (-SAHA, C) for 48 hours, as well as an isotype control (A). The Table (B) shows the percentage of cells lying within the gated area, containing Zta positive cells, as based on the isotype mouse control.

Based on RNA levels, western blot results and previous, unpublished findings by the Tsao laboratory, an increase in EBV genome load in induced cells of approximately 15 fold would be expected. However, qPCR analysis using primers specific for either the human or EBV genome did not lead to the expected results (Figure 14) and no significant difference of EBV load could be detected between

induced and uninduced cells. The amount of EBV per host cells was shown to be approximately equal to one.

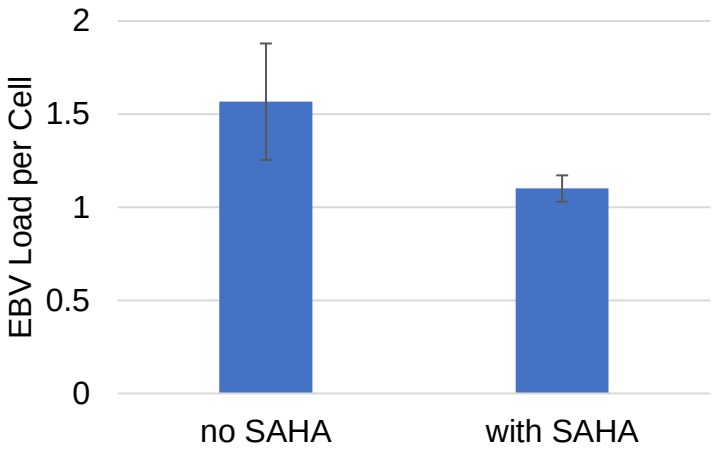


Figure 14: qPCR results showing the amount of EBV relative to the amount of human genome. EBV lytic cycle was induced in HONE1-EBV cells with 10μM SAHA (with SAHA) or uninduced (no SAHA) for 48 hours. The EBV genome was detected using primers specific for EBV polymerase. Human genome was detected using primers specific for globin. Values are represented as mean ± SD. There is no significant difference between samples treated with SAHA in comparison to untreated samples ($p<0.05$).

Two distinct populations of HONE1-EBV cells can be observed, one attached to the flask, as would be expected of an epithelial cell line, and one freely floating in the medium. The previous results have been based on the entire cell population. Through casual observation, it has become clear that the ratio of floating:adherant cells increases when SAHA is added, to a point where at least 50% of the cells are detached from the flask after 48 hours. To determine whether both cell populations show evidence of lytic induction, a time course was set up

at two different concentrations of SAHA and the presence of the early lytic protein Zta was determined. The overall EBV load per cell was also assessed.

As shown in Figure 15A, SAHA successfully induces the expression of Zta in adherent cells.

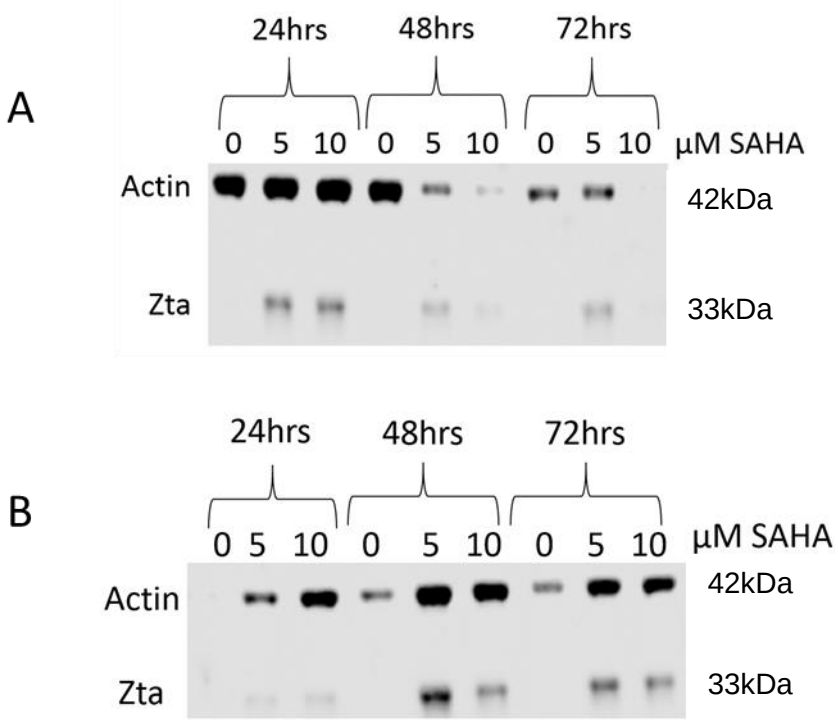


Figure 15: Western Blots showing the expression of Zta and actin in adherent or floating HONE1-EBV cells. Lytic cycle was induced with 0µM, 5µM or 10µM SAHA at 24hrs, 48hrs and 72hrs post induction. An antibody specific for actin was used as a loading control. **A**: Adherent cells, **B**: Floating cells.

Figure 15B shows a western blot of actin and Zta levels in floating cells which demonstrates that Zta levels are also increasing in this population when cells are induced with SAHA. As opposed to adherent cell numbers, the numbers of floating cells increase with the addition of SAHA, as measured by actin levels.

Relative to actin levels, the effect of SAHA on Zta expression, and thus early lytic cycle, seems to be strongest at time points past 24 hours using 5μM or 10μM SAHA. The amount of Zta detected in floating cells is shown to be slightly lower when compared to adherent cells.

As part of the same experiment, the genome copy number of EBV per host cell was determined at different concentrations of SAHA and an induction time of 24, 48 or 72 hours in floating and adherent cells. Results are summarised in Figure 17. The floating cell population shows no significant increase in EBV copy number under the influence of SAHA. Adherent cells show a clear increase in EBV copy numbers when induced with both 5μM and 10μM SAHA over time, with significantly more EBV present by 72 and 48 hours in floating than adherent cells.

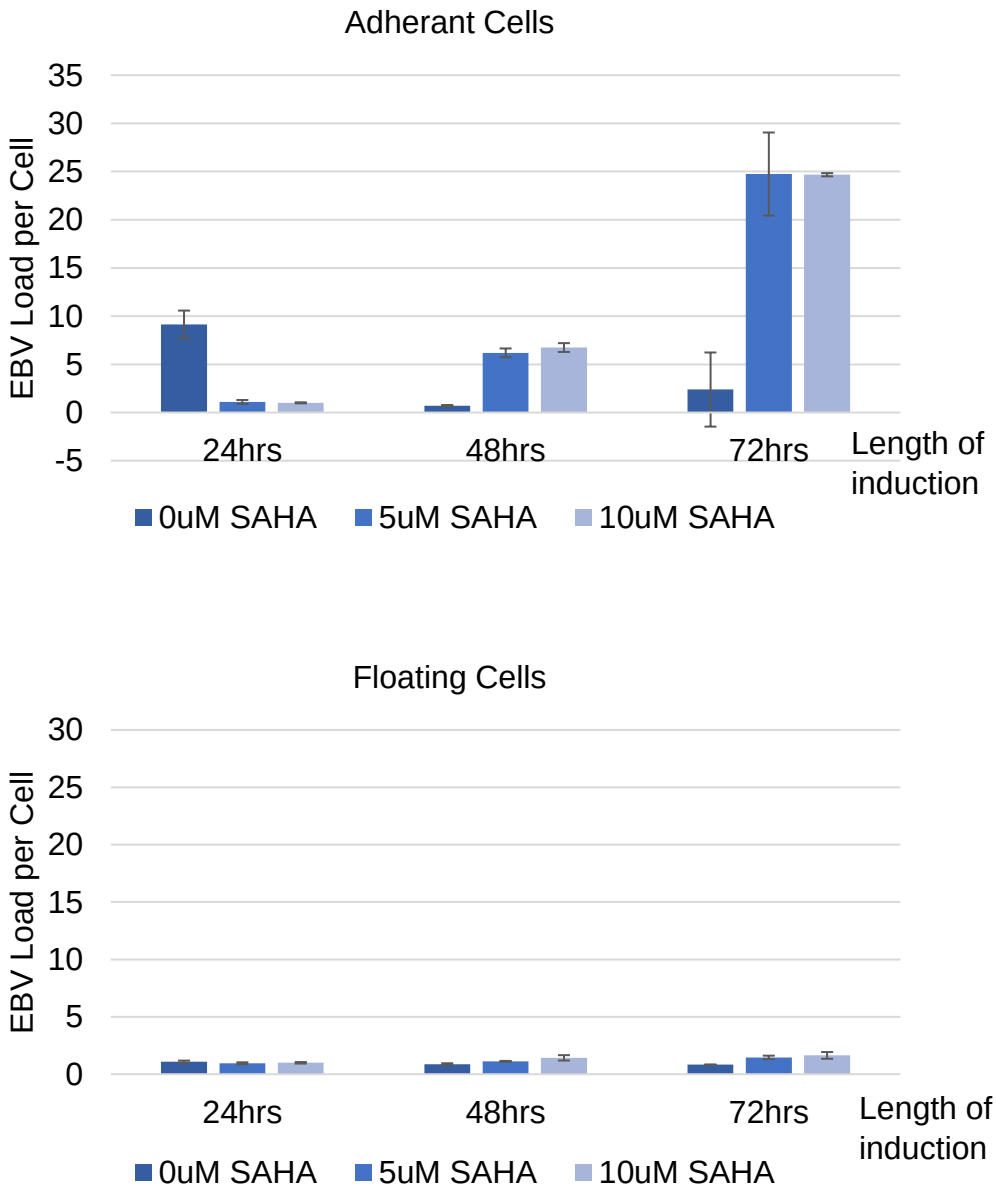


Figure 17: qPCR results showing the relative amount of EBV genome copies per human genome numbers in adherent and floating HONE1-EBV cells. Lytic cycle was induced with 0μM, 5μM or 10μM SAHA with 24, 48 and 72 hours of induction. This experiment shows a significant increase of EBV load in cells treated with 5μM or 10μM SAHA in comparison to untreated cells at 48hrs as well as samples treated with 10μM SAHA at 72hrs ($p<0.05$). Averages are based on duplicate pipetting repeats. Values are represented as mean ± SD.

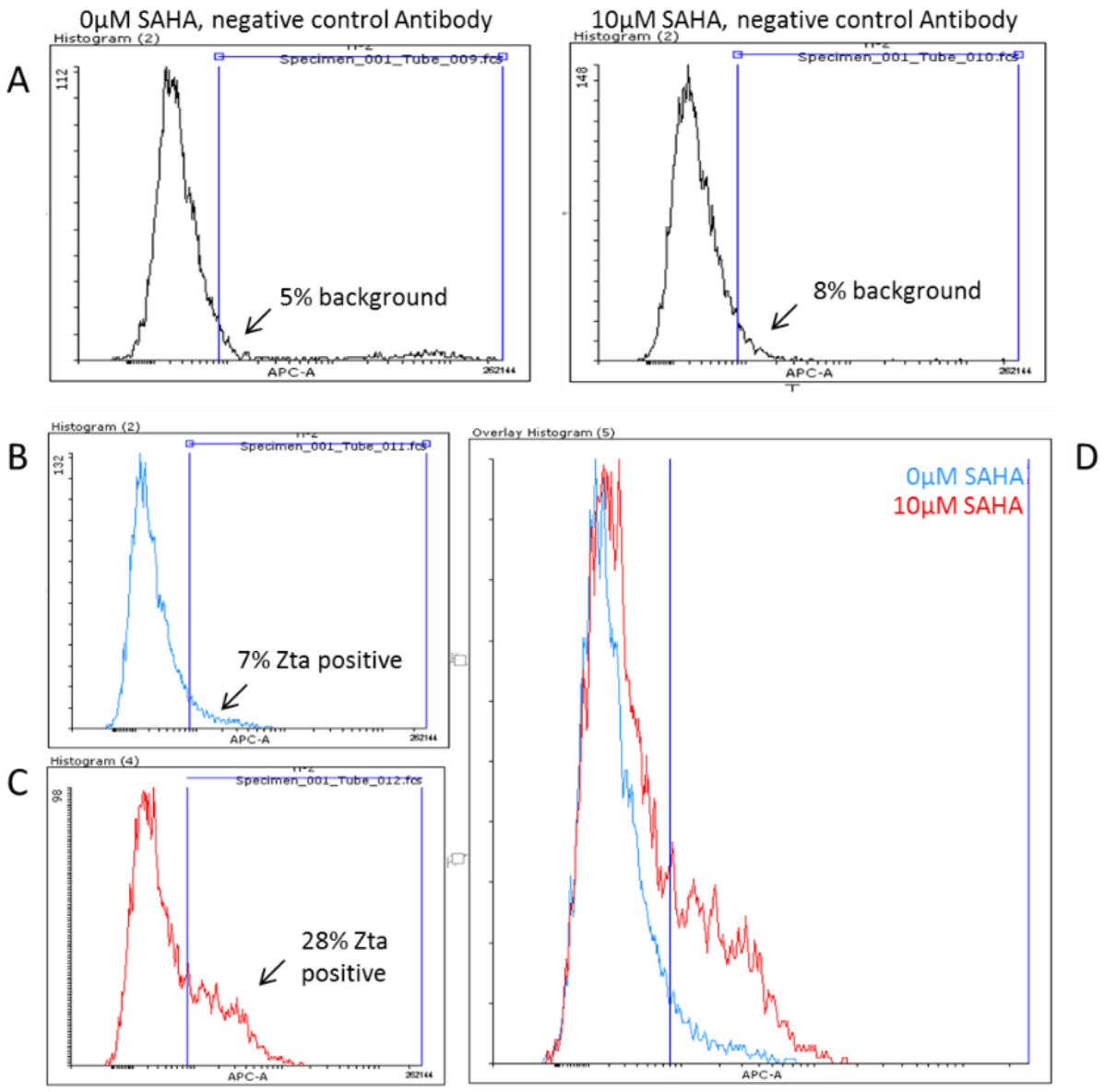


Figure 18: Histograms showing the percentage of HONE1-EBV cells containing Zta. Adherent HONE1-EBV cells were induced with 10µM SAHA for 48 hours, or not induced, and stained for Zta and a negative control antibody. A: Induced and un-induced cells stained with a negative control antibody showing background fluorescence. B: Un-induced cells stained for the detection of Zta levels in the cell population. C: Induced cells stained for the detection of Zta levels in the cell population. D: Overlay Histogram of un-induced and induced cells stained for Zta.

Based on these results, FACS was performed to detect the percentage of adherent HONE1-EBV cells positive for Zta. Results are summarised in histograms shown in Figure 18. Staining with the control antibody shows a relatively high background of 5-8%. As before, Zta positive cells lie within the gated region of the histogram, which accounts for 7% of uninduced cells (0μM SAHA) and 28% of induced cells (10μM SAHA). Therefore approximately 21% of the cell population contain EBV which has entered into lytic cycle following induction, representing a 3.7 fold increase in Zta levels in comparison to un-induced cells. These FACS results confirm a clear difference in Zta expression when EBV lytic cycle is induced in adherent HONE1-EBV cells with 10μM SAHA for 48hrs in comparison to untreated cells.

Consequently, the following results will be based on the adherent HONE1-EBV cell population only.

3.3 Chromatin Immunoprecipitation (ChIP)

In this section the procedure for preparing chromatin from HONE1-EBV cells for Zta ChIP qPCR and Zta ChIP sequencing is outlined. Different amounts of sonication were tested and the sheared chromatin was examined for the presence of Zta using western blotting. ChIP qPCR was performed to investigate the enrichment of Zta binding regions to the viral genome, which were previously mapped to the viral genome in B cells [18] (BCL2A1 data not published).

Chromatin was prepared from HONE1-EBV cells treated with 10μM SAHA for 48 hours. To optimize the sonication of this chromatin, a range of time point samples

was taken and the samples were separated on a DNA gel to determine their size (Figure 19). As the desired size for subsequent ChIP sequencing library preparation is 200bp, the ideal chromatin size for efficient ChIP analysis lies around this size region. While we would expect to see a narrower size range of chromatin fractions, it is clear that some of the chromatin lies within this area. 10 rounds of sonication were used for chromatin preparation.

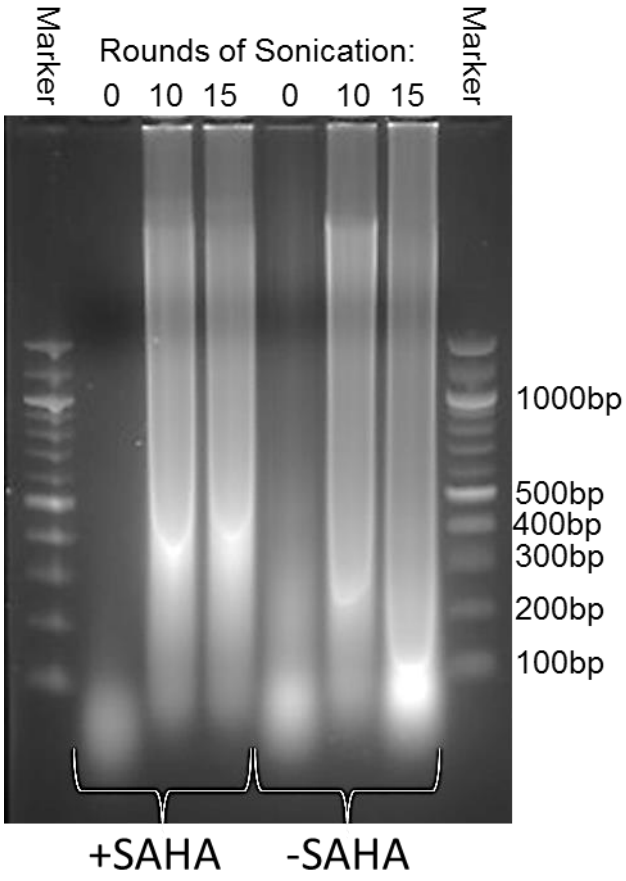


Figure 19: Agarose Gel of chromatin samples extracted from HONE1-EBV cells. EBV lytic cycle was induced with 10μM SAHA (+SAHA) or uninduced (-SAHA), after 0, 10 or 15 rounds of sonication in a volume of 1.5ml at 30% amplitude with a pulsar of 10sec on, 10sec off.

To demonstrate that lytic cycle had been induced in the HONE1-EBV cells when chromatin was extracted, the sample was reverse cross-linked and western blot analysis was undertaken to determine whether Zta is present in the induced sample.

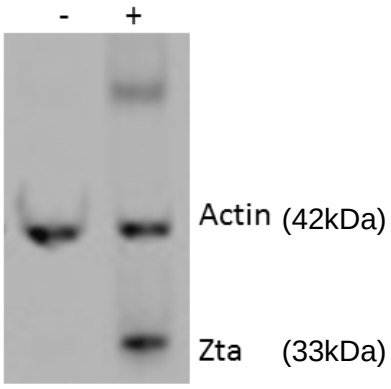


Figure 20: Western blot showing the expression of Zta in chromatin from HONE1-EBV cells. Lytic cycle was induced (+) or uninduced (-) with 10μM SAHA for 48 hours. An antibody specific for actin was used as a loading control.

The western blot shown in Figure 20 demonstrates that Zta is expressed in induced cells. In addition, a second band has been detected by the Zta specific antibody. This band is approximately twice the size of the Zta protein and is thought to show Zta dimers which have not been completely reverse cross-linked.

A chromatin immunoprecipitation (ChIP) assay was carried out on the sonicated chromatin to assess whether Zta binds to the epithelial cell and viral genome in HONE1-EBV cells. For this purpose, primers specific for regions of the host cell genome (BCL2A1) and the viral genome (OriLyt), to which Zta has been shown to bind in Akata B cells, were used. Results are summarised in Figure 21 and clearly show enhanced binding of Zta to the origin of lytic replication (OriLyt) on

the viral genome in comparison to an OriLyt flanking region. This flanking region is used as a negative control together with an IgG isotype control. A fold change of 2.4 can be identified between Zta binding to OriLyt in induced versus uninduced cells.

In addition to the origin of lytic replication, Zta association with a second B cell Zta binding region, the B-cell lymphoma 2-related protein A1 gene (BCL2A1), and it's flanking region, was also determined. A significant increase with a fold change of 8.5 in Zta binding to this region of the epithelial cell genome can be seen when lytic cycle is induced.

The ChIP protocol, which was kindly provided by Rajaei Almohammed, states a working range of ChIP antibody between 2µg and 10µg. Both extremes of Zta specific antibody were used for this ChIP to investigate which amount results in the most efficient ChIP. Based on the results shown in Figure 21, 10µg of Zta antibody was used for future ChIP experiments.

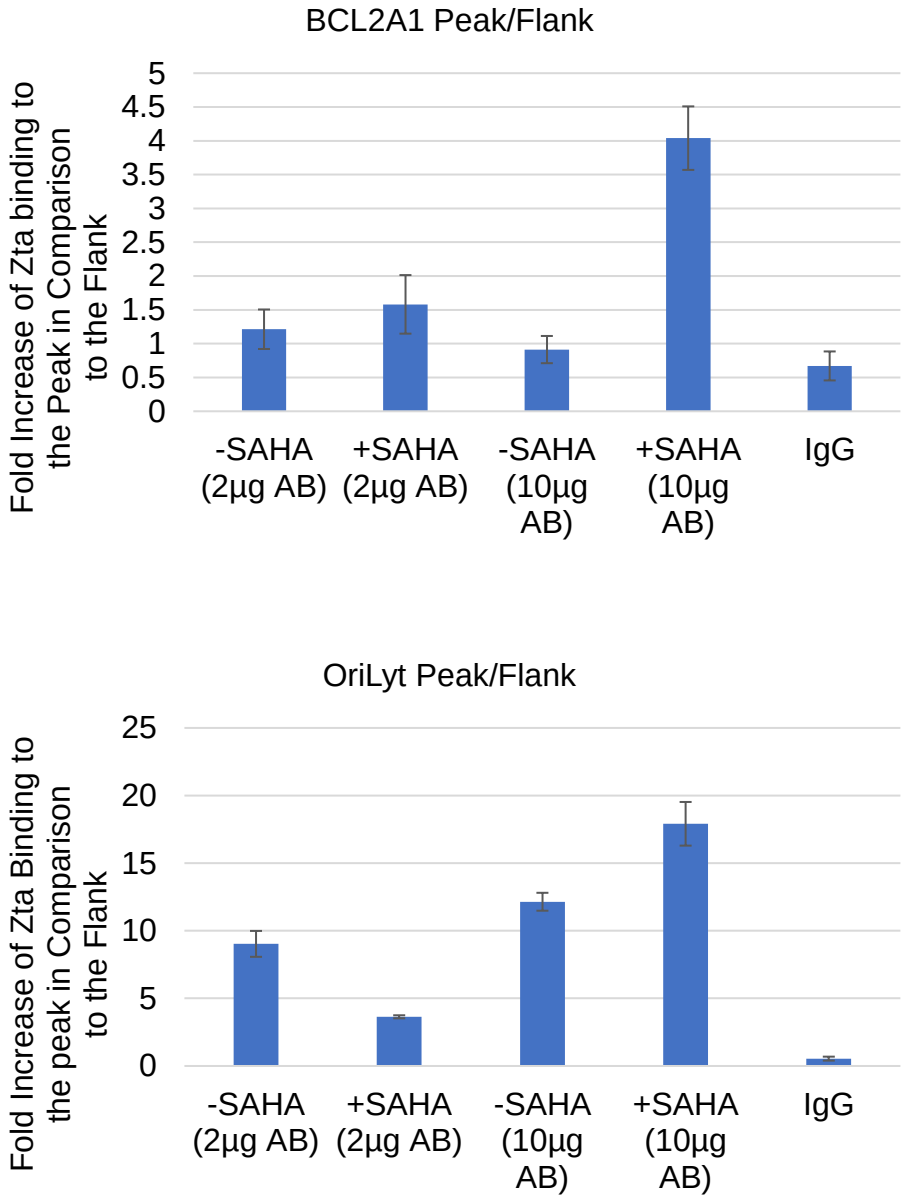


Figure 21: qPCR results of ChIP experiment showing Zta binding to BCL2A1 or OriLyt in comparison to flanking regions and an IgG control. HONE1-EBV cell were induced with 10µM SAHA (+SAHA) or uninduced (-SAHA) for 48 hours. ChIP was performed using either 2µl or 10µl of BZ1 antibody (200µg/ml). The difference of Zta binding to the peak vs flanking region in cells treated with SAHA is significant ($p<0.05$). Values are represented as mean ± SD.

3.4 Library Preparation and Quality Control

Based on the results outlined in the previous section, the role of Zta in epithelial cells was examined more closely by coupling chromatin immunoprecipitation to next generation sequencing (ChIP Seq). For this purpose, the ChIP assay was scaled up in duplicates and the DNA concentration in the sample was determined using a NanoDrop. The samples were prepared for sequencing (library preparation) by end repair, dA-tailing, adaptor ligation and PCR enrichment, followed by size selection of the library.

qPCR analysis was carried out on the large scale ChIP DNA samples to verify the presence of DNA in the samples going into the library preparation. This confirmed that DNA was present and showed enriched binding of Zta to OriLyt in comparison to a flanking region in both ChIP assays (Figure 22). However, a problem was encountered when the DNA could not be detected during the size selection step. Attempts to re-stain the gel with SYBRSafe did not improve the detection of the sample, even on an over-exposed gel (Figure not shown). Further investigation of the problem finally revealed an error during the PCR enrichment step involving the absence of a required enzyme which is required for the adaptor used during this experiment that contains a modified nucleotide requiring cleavage by this enzyme before PCR amplification can occur.

As enough ChIP DNA was still available, a second library preparation was attempted. The concentration of the DNA was determined on a Qubit, which is more sensitive than the NanoDrop used previously. The recommended 10ng DNA was used to prepare the library. In addition, a higher amount of 90ng of DNA was used as starting material for the library preparation. The concentration of the end repaired, dA-tailed, adaptor ligated and PCR enriched DNA was determined

before loading it onto the size selection gel to avoid loss of library samples. In addition, the DNA stain SYBRGold was used to be able to detect a smaller amount of DNA. The appropriate sized samples (175-225bp) were excised from the TAE gel and purified. qPCR analysis of the prepared sample was carried out, showing a clear enrichment of DNA bound by Zta in the ChIP samples compared to the input control (Figure 23). The fold enrichment for the 90ng samples is equal to 1.16 for the Input, 13.64 for ChIP1 and 8.39 for ChIP2. The fold enrichment for the 10ng sample cannot be calculated as OriLyt Flank readings were too low to be detected during this qPCR step. All six samples were sent to an external laboratory (Euorofins Genomics group) for high-throughput sequencing.

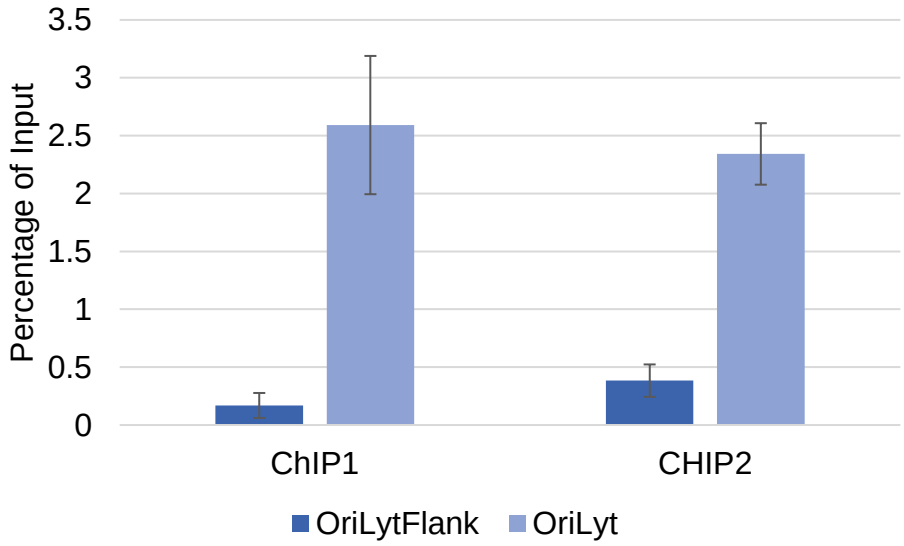
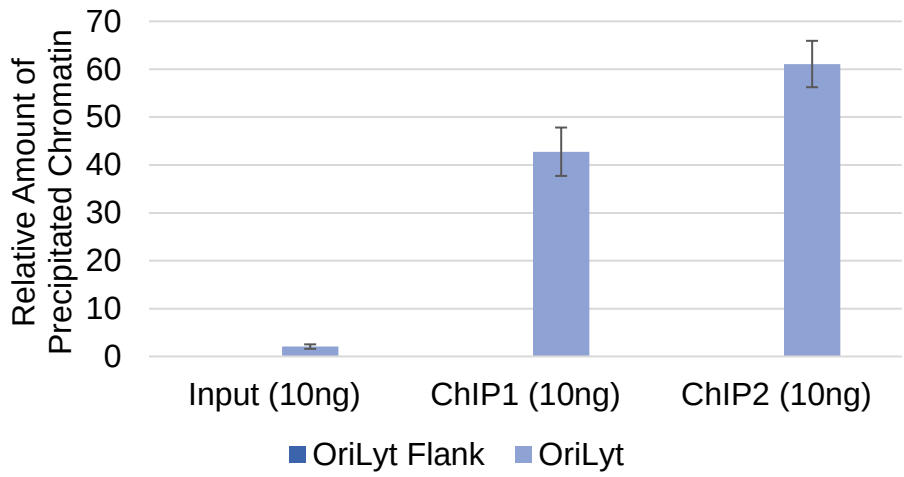


Figure 22: qPCR results of large scale ChIP experiment. Graph shows significantly enhanced Zta binding to OriLyt in comparison to flanking regions ($p<0.05$) in duplicate ChIPs. HONE1-EBV cell were induced with 10μM SAHA (+SAHA) or uninduced (-SAHA) for 48 hours. Values are represented as mean ±



SD.

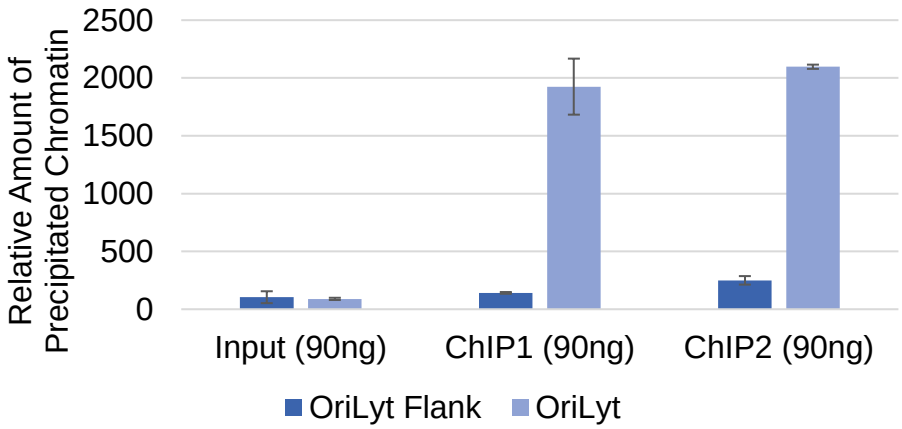


Figure 23: qPCR results showing the relative amount of OriLyt and OriLyt Flank DNA bound by Zta in a HONE1-EBV ChIP sequencing library. Lytic cycle was induced with 10μM SAHA for 48 hours or uninduced. 10ng (A) or 90ng (B) ChIP DNA were used to prepare the library. The aount of DNA bound by Zta in the flanking region of OriLyt in the 10ng ChIP was too low to be detected in the qPCR experiment. The amount of DNA bound by Zta is relative to a standard curve where 90ng input was assigned the value of 100. The amount of binding of Zta to OriLyt is significantly larger than to the flanking region in both 90ng ChIPs ($p<0.05$). OriLyt Flank readings for the 10ng ChIP were too low to be detected during qPCR. The Values are represented as mean ± SD.

To determine the quality of the ChIP library, the sample was evaluated with the help of a Bioanalyser by the Euorofins Genomics group. This makes it possible to determine the DNA fragment sizes within the samples and the corresponding concentrations. The resulting electropherogram (Figure 24) shows the sizes and concentrations of the DNA fragments in the library with the far left peak, or bottom band (green), and far right peak, or top band (purple), representing markers.

Ideally, no adaptor should be present to contaminate the library sample, and the concentration of the immunoprecipitated DNA should be ≥10nM. The actual concentration of this library has been shown to be 1.6nM (244pg/µl). Hence the library could not be sequenced.

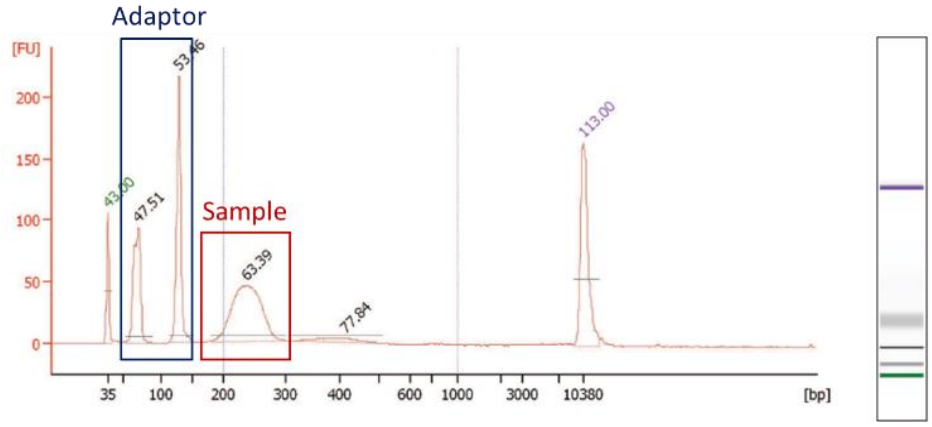


Figure 24: High-Sensitivity Electropherogram of library sample prepared for sequencing. Diagram representative of ChIP1, ChIP2 and input sample. The blue box indicated the size region of the adaptor and the red box represents the desired size, in base pairs, of the sample that is to be sequenced. Numbers above peaks represent the time, in seconds, at which the concentration at the particular size was measured.

As a result of the failed library quality control, an additional large-scale chromatin preparation and ChIP experiment were performed to construct a library. To investigate if the chromatin was sonicated sufficiently a southern blot was performed which is shown in Figure 25. In addition, a ChIP qPCR experiment was performed to demonstrate if the lytic cycle was successfully induced. As is shown in Figure 26, there is a significant enrichment of Zta binding to OriLyt in comparison to a negative control region (OriLyt Flank) suggesting that the chromatin fragments have been immunoprecipitated successfully. Following these results a library was prepared for sequencing using the recommended 10ng ChIP DNA.

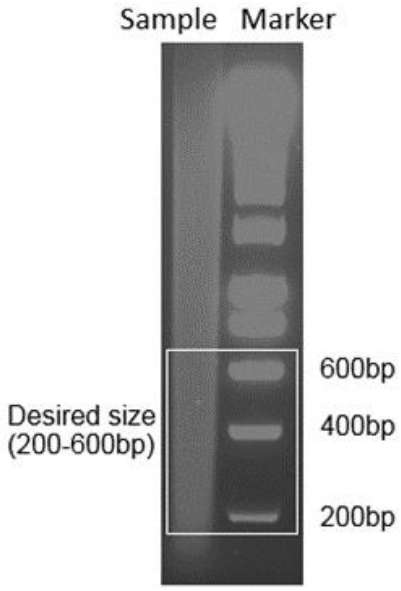


Figure 25: Agarose Gel of chromatin samples extracted from HONE1-EBV cells. Lytic cycle was induced with 10μM SAHA, after 36 rounds of sonication in a volume of 1.5ml at 40% amplitude with a pulsar of 10sec on, 10sec off.

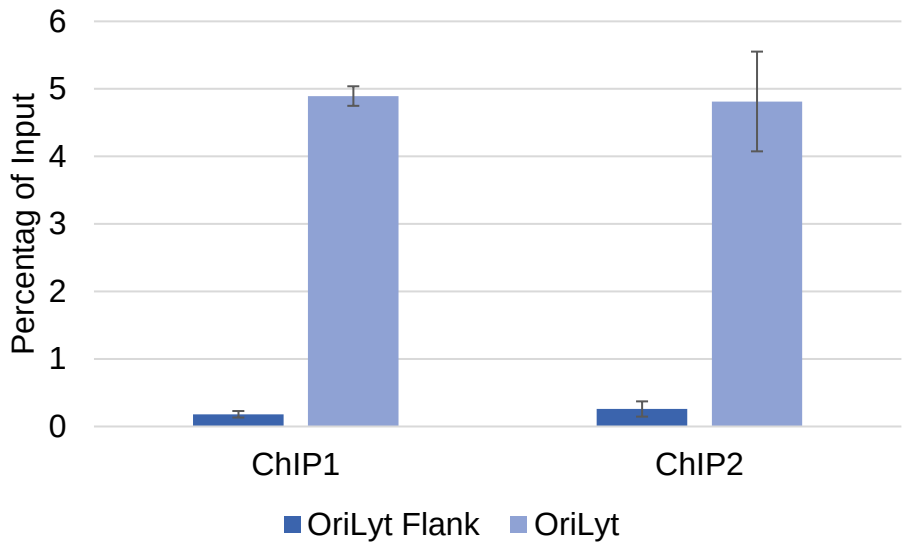


Figure 26: ChIP qPCR analysis showing the relative amount of OriLyt DNA bound by Zta in HONE1-EBV cells. Lytic cycle was induced with 10μM SAHA for 48 hours. The difference in binding of Zta to OriLyt and OriLyt Flank is significant in both ChIP experiments ($p<0.05$). Data shown as percentage of Input. Values are represented as mean ± SD.

To determine the quality of this library, the samples were evaluated on a Bioanalyser before sending them for sequencing. Results are shown in Figure 27.

No ChIP sample could be detected during this Bioanalyzer run. The high-sensitivity kit used previously is able to detect concentrations of as low as 100 pg/μL DNA. The kit used during this run should be able to detect the required 10nM concentration, as can be seen for the Input sample. However, if the sample concentrations are below the 0.5ng/μl detection limit of this kit, they would not be identified. This is the case for the ChIP samples. In comparison to the previous quality control shown in Figure 24, no adaptor contamination can be detected.

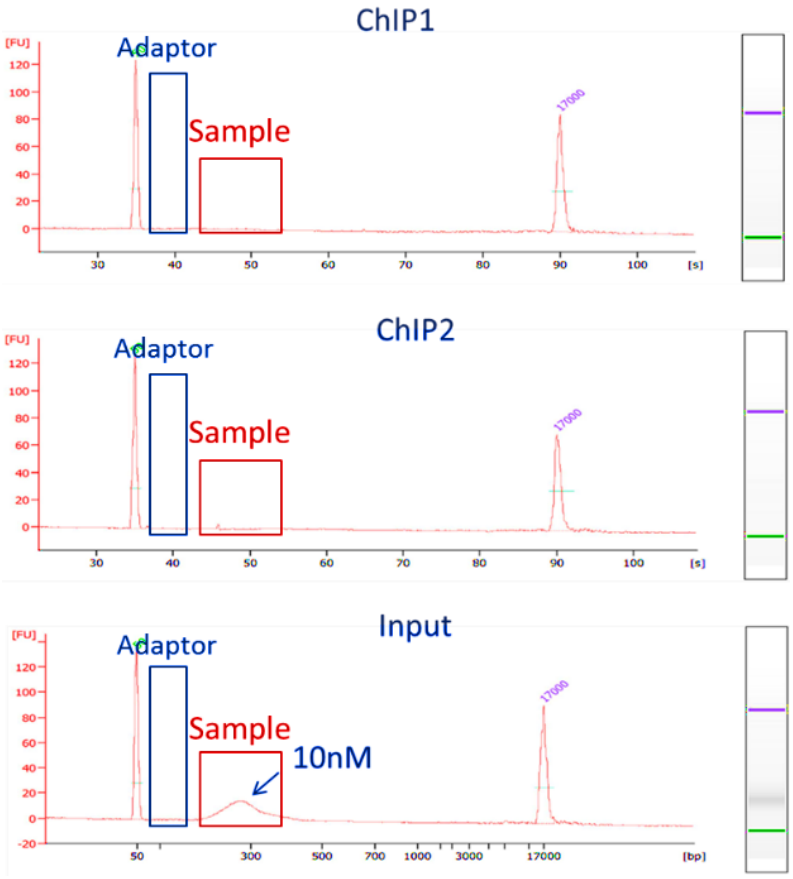


Figure 27: Electropherogram of library samples prepared for sequencing. As before, the blue box indicates the size region where the adaptor would be present and the red box represents the desired size, in base pairs, of the sample that is to be sequenced.

The low sample concentration at the selected size range highlights the need for chromatin sonication to produce as high a concentration of 200bp fragments as possible. For this purpose the sonication protocol was modified to achieve a more favourable fragment size for ChIP sequencing purposes (Material and Methods “Chromatin Sonication II”). In a nutshell, cells were 15x more concentrated and the crosslinking time was reduced from 15 to 10 minutes to prevent over-

crosslinking. The samples were sonicated in a 15x lower volume (100µl instead of 1.5ml) and the sonication amplitude was lowered from 40% to 21%. In addition, samples were vortexed after every three cycles to ensure uniform sonication of the whole sample.

The ideal fragment size for sequencing ChIP DNA is 200bp. Based on the results shown in Figure 28, the new chromatin lot was prepared using the improved chromatin sonication protocol.

The Zta ChIP was performed as before and immunoprecipitated DNA was assessed using qPCR (Figure 29).

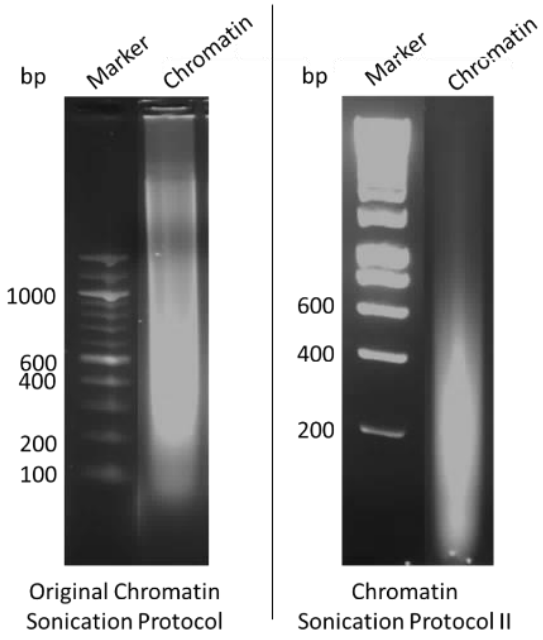


Figure 28: Agarose Gel showing the size, in base pairs (bp), of sonicated HONE1-EBV chromatin fragments. Lytic cycle was induced with 48µM SAHA for 48 hours, using the original chromatin sonication protocol (left) in comparison to the altered protocol (right).

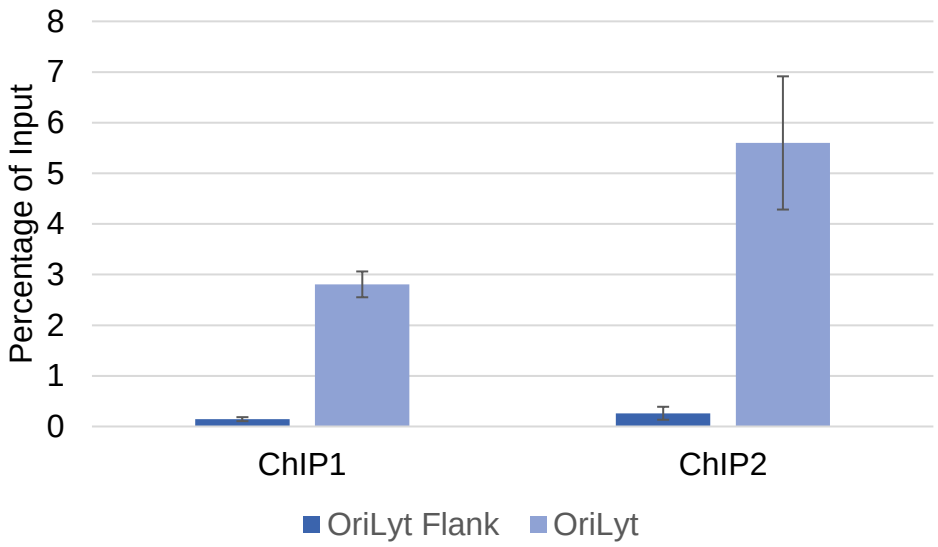


Figure 29: ChIP qPCR analysis showing the amount of viral OriLyt DNA or flanking region bound by Zta in new HONE1-EBV chromatin. Lytic cycle was induced with 10μM SAHA for 48 hours. OriLyt represents the origin of lytic replication on the EBV genome, which is known to be bound by Zta. OriLyt Flank represents a negative control region in the vicinity of OriLyt which is not bound by Zta. The enhanced binding of Zta to OriLyt in comparison to the flanking region is significant ($p<0.05$). Data shown as mean ± SD.

As shown in Figure 29, the chromatin was enriched for DNA specifically bound by Zta, as desired, and a library was constructed to prepare the DNA for sequencing. After the library preparation, the samples were analysed with the help of a high-sensitivity Bioanalyser assay. The results of this quality control step can be seen in Figure 30 below, which shows 8-15x higher sample concentrations at the desired size than required (>10nM), with no adaptor contamination.

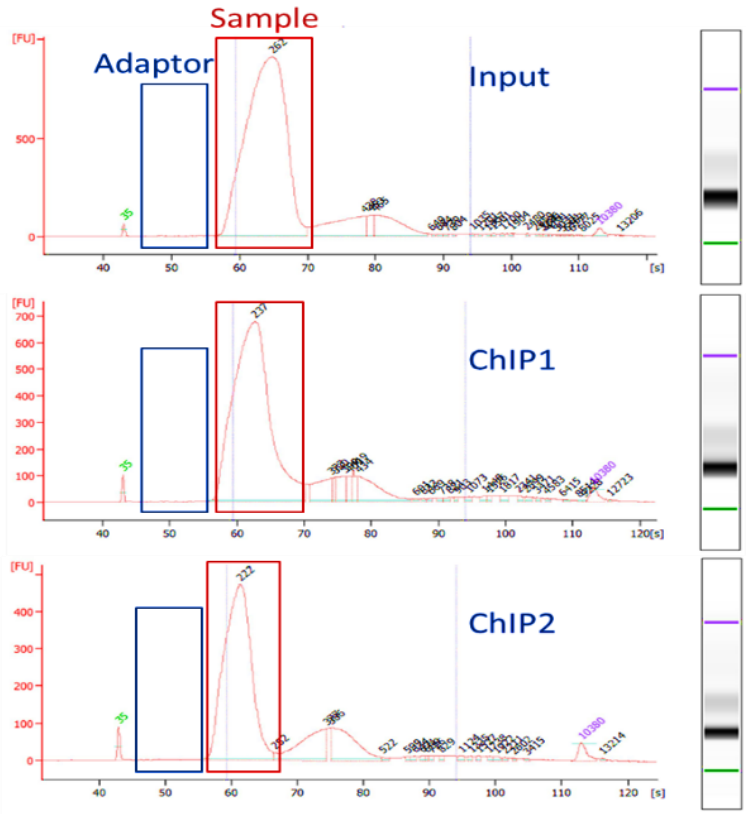


Figure 30: High-Sensitivity Electropherogram of library samples sent for sequencing. As before, the blue box indicates the size region where the adaptor would be present and the red box represents the desired size, in base pairs, of the sample that is to be sequenced. Input Sample peak: 151nM, ChIP1 Sample peak: 81nM, ChIP2 Sample peak: 48nM.

Figure 31 shows a number of different library preparations on agarose gels, which have been used during the size selection step of the library preparation process, and summarises the various stages and alterations made to achieve an effective ready-to-sequence library.

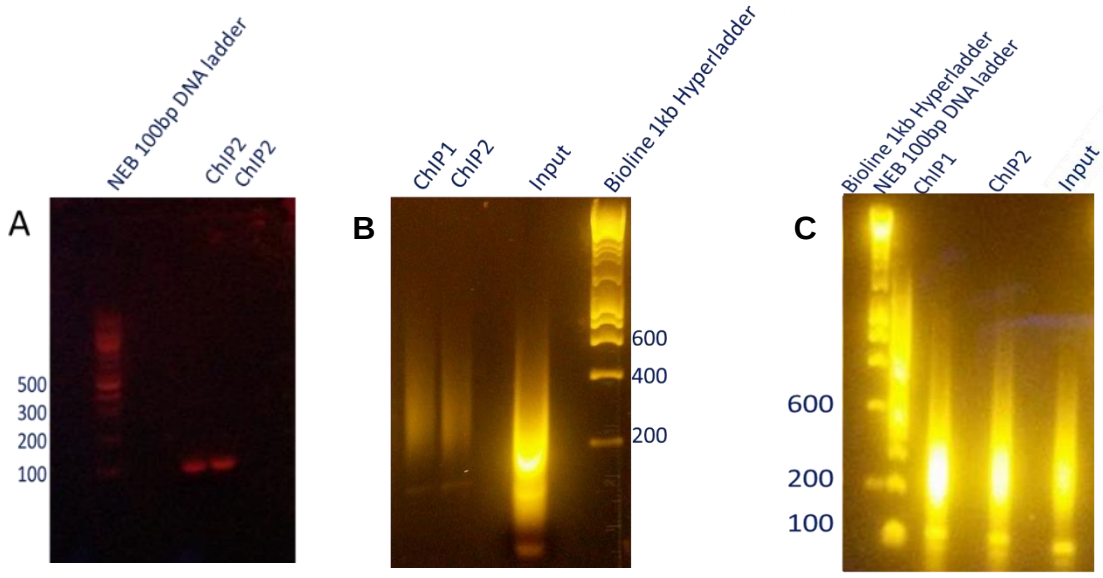


Figure 31: Size selection gels showing various library preparations. Red boxes represent the region of the gel selected for purification and consequent sequencing. A: First attempt to prepare a ready-to-sequence library using SYBR Safe dye to visualise the DNA. This blot shows one of the duplicate ChIPs only (ChIP2), but is representative of all samples. No DNA could be detected at the desired size as an essential enzyme had been omitted from the amplification step of the library preparation process. B: This blot is representative of the next two library preparation attempts. SYBR Gold dye was used to visualise DNA. Sample concentrations were too low to be sequenced. C: Size selection gel of a successful ready-to-sequence library using the improved sonication protocol to prepare chromatin.

3.5 Zta binding to the viral genome

Following the sequencing of the ChIP libraries by an external laboratory, 30.8GB of data was received in *.bam* (binary) format. This data was processed using a range of data analysis platforms.

Zta ChIP sequencing enrichment was identified for several sites on the EBV genome. Figure 32 shows an example of high enrichment for a site around the 42kb region of the EBV genome which is in the area of the BHLF1 gene promoter (OriLyt).

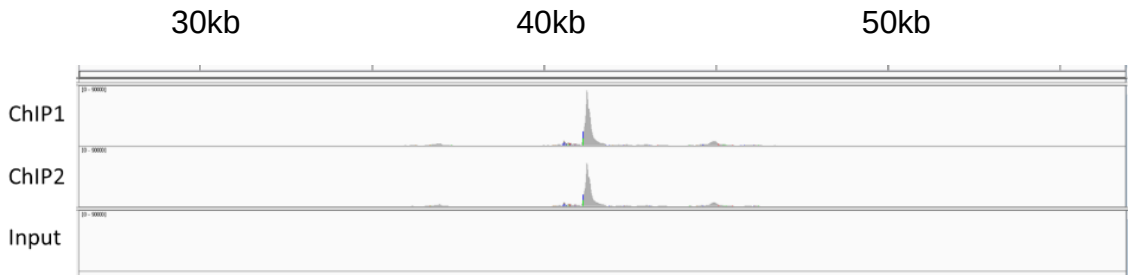


Figure 32: Example of Zta binding to the EBV genome in HONE1-EBV cells visualised through the display of Zta ChIP sequencing data on the Integrative Genomics Viewer (IGV). Data range 0-90000 reads. Figure shows an example of high enrichment for a site around the 42kb region of the EBV genome which is in the area of the BHLF1 gene promoter (OriLyt) that is associated with the EBV lytic cycle.

The large majority of Zta ChIP sequencing analysis focussed on the human genome. This is discussed in more detail in the following chapter.

3.6 Discussion

Results show that Zta and BLLF3 are upregulated at the RNA level in HONE1-EBV cells under the influence of SAHA. In addition, it has been confirmed by western blotting that Zta is upregulated at the protein level, indicating that lytic cycle is successfully being induced. However, FACS results initially did not

support this conclusion. It is possible that induction was inefficient due to a range of potential causes, including unequal distribution of the SAHA in the defrosted stock, an inexact induction time or cell density and number in the flask at the time of induction. It is also possible that the antibodies used was not suitable to detect Zta expression in the cell line. Additionally, it is likely that only a minority of HONE1-EBV cells are in lytic cycle. FACS analysis takes every cell into account individually which may explain part of the obtained result.

Detection of the EBV load in induced cells at first did not show the increase expected during lytic replication. These findings may indicate that the virus is at a very early stage of the lytic cycle. It is possible that, at 48 hours of induction, only a fraction of the cells have entered productive lytic cycle.

Subsequent to these initial results, the induction conditions were optimised to achieve more efficient results for investigating the sites on the host cell and viral genome bound by Zta in epithelial cells.

The amount of actin present in the sample is proportional to the number of cells, which leads to the conclusion that the number of adherent HONE1-EBV cells declines with increasing exposure to SAHA. It is possible that the floating cells are apoptotic due to either the direct effects of SAHA on the cell, the loss of the EBV plasmid (which is selected for with G418), or the completed lytic cycle. All six samples of floating cells which were treated with SAHA show no evidence of EBV replication. This may be due to the absence of lytic cycle, or its completion, in these cells. Alternatively it is possible that the cells detach from the flask due to differentiation, which has been suggested to be associated with lytic replication of EBV in epithelial cells [130].

The determination of average EBV genome number per cell revealed a significant lower amount of EBV in the floating compared to the adherent population. At 24 hours the genome copy number of EBV seems higher in cells not containing SAHA which could be due to the spontaneous entry into lytic cycle, as HONE1-EBV cells are not tightly latent. Based on these results, it was decided to work with adherent cells only in following experiments. As we are interested in the early lytic cycle and cell numbers decrease with time, it was decided to use cells induced for 48 hours with 10μM SAHA.

Based on FACS results of HONE1-EBV cells induced in this way, approximately 21% of the HONE1-EBV cell population have entered into lytic cycle following induction. Hence, any effects observed as a result of this induction, during experiments based on the average of the whole population, actually only represent less than a quarter of cells. These experiments include the EBV copy number per cell, RNA levels, and western blots. Zta ChIP qPCR experiments are not included in this, as the immunoprecipitation of the early lytic cycle protein, Zta, selects cells in lytic cycle. However, it does point to the necessity to increase the number of cells used for Zta ChIP experiments in these cells, in comparison to other cell types which can be induced more efficiently.

It was shown that Zta binds to both the viral and host cell genome in HONE1-EBV cells when EBV is induced into lytic cycle using SAHA. Based on these results, chromatin was prepared and Zta binding sites were immunoprecipitated. A ChIP sequencing library was prepared.

The first Zta ChIP ready-to-sequence library quality controls had a too low sample concentration to be sufficient for sequencing. It is thought that this is due to a

chromatin preparation process of low quality. Adaptor contamination is most likely a result of poor size selection. Alternatively it could be a contamination of the sample in the size range of the adaptor. It was clear that these processes needed to be improved to achieve the desired results. Subsequent changes made to the chromatin preparation and sonication did indeed yield a more favourable result for ChIP sequencing purposes as the following Zta ChIP and library preparation generated over eight times more sample than required, with no adaptor contamination. After the optimisation of this process, as outlined in the previous section, the samples passed quality control and were sent to an external laboratory for high-throughput sequencing.

The analysis of Zta ChIP sequencing reads mapped to the EBV genome revealed several clear Zta binding sites. Further analysis is necessary to reveal any differences and similarities of Zta binding to the viral genome in epithelial cells compared to B cells. Following up on these results could provide insights into the process of the EBV lytic cycle activation in epithelial cells, which would lead to a better understanding of the biology of EBV infection.

Analysis of Zta binding to the human genome is presented in the following chapter.

Chapter 4: Zta ChIP sequencing data analysis of binding to the human genome

4.1 Introduction

The previous chapter outlined how a HONE1-EBV Zta Chromatin Immunoprecipitation (ChIP) Library was prepared, in duplicate, and sent to an external laboratory for sequencing and mapping to the host cell and EBV genome. This resulted in 30.8GB of data which was received in *.bam* (binary) format. This chapter will focus on the analysis of the data to identify Zta binding sites on the host genome, using a range of data analysis platforms. The first steps in analysing the ChIP sequencing data were performed with the help of deepTools Galaxy which includes a public server that contains a range of python tools which make it easier for researchers without a background in informatics to analyse high-throughput sequencing data, like ChIP sequencing data [116].

The resulting output of deepTools Galaxy enabled the use of the Genomic Regions Enrichment of Annotations Tool (GREAT) to analyse the functional significance of distal and proximal cis-regulatory elements which have been identified by Zta binding across the human genome [120].

Gene lists obtained through GREAT were further analysed using EnrichR, which is a gene list enrichment analysis tool. The list of genes associated with Zta binding across the human genome was used as input for computing enrichment. EnrichR is based on existing lists created from prior knowledge which has been organized into gene-set libraries. [119, 131].

Furthermore, DNA motifs of Zta binding peak regions were discovered, analysed and visualised using the online tool MEME-ChIP [132].

Analysis revealed a range of host genes that are bound by Zta in epithelial cells. In addition comparisons of Zta binding sites between epithelial and B cells were undertaken. Histone modification and transcription factor binding in various cell types were also investigated to identify common factors. For this purpose, the Encyclopaedia of DNA Elements (ENCODE) database was utilised [133]. Results of this data analysis were followed up with further experiments.

4.2 Mapped sequencing *.bam* file data analysis

To analyse the semi-compressed alignment files (*.bam*) received from the Zta ChIP sequencing experiment, reads mapped to the human genome were transferred to deepTools Galaxy using a file transfer protocol (FTP). To assess the enrichment strength of the duplicate Zta ChIPs, a bamFingerprint graph was produced (Figure 33). This shows that approximately 18% of genomic bins contain the greatest fraction of reads, which is not considered to be very strong enrichment. The bamFingerprint graph created by our data is more characteristic of a broad, histone modification ChIP rather than high, narrow enrichment that would be expected for a transcription factor, like Zta [134].

As we have checked the chromatin fragment size prior to ChIP sequencing, it is unlikely that the fragments were so long as to impair enrichment resolution and specificity. Consequently we investigated the possibility that many binding sites are close to each other, which would lead to a large fraction of bins having a high coverage.

Additionally, even the input file contains a large amount of bins with zero ready as indicated on the x-axis. It would be expected that the input does not map completely to the human genome. Due to the relatively short fragment size of around 200bp, the unique mapping to highly repetitive regions of the genome is not possible. Additionally some of the DNA present in the input sample will be of EBV origin. Even so, 40% of the genome containing no reads is unusual. As the origin of HONE1-EBV cells used in this experiment has not been identified, it is possible that some human genome was lost or that a cell line from a different species was involved in the establishment of this cell line.

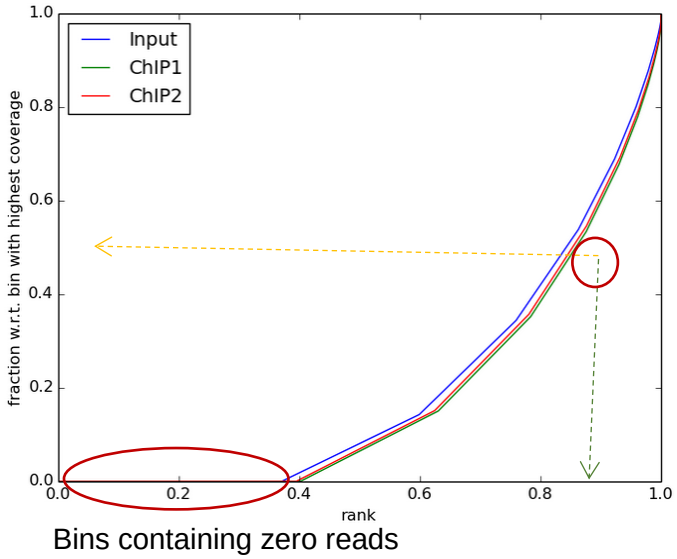


Figure 33: bam Fingerprint graph showing the enrichment of the duplicate HONE1-EBV Zta ChIPs and input. Cells were induced with 10μM SAHA for 48 hours. The cumulative sum of read counts was plotted against bin read-count rank. When counting the reads contained in ~82% of all genomic bins, ~50% of the maximum number of reads are reached ~18% of the genome contain the greatest fraction of reads. Fragment: 200bp, Bin size: 30bp, Bins ranked: highest number of reads, highest rank.

Figure 34 shows an example of binding sites clustering around the FOSB gene on the human genome. This binding pattern of Zta explains the unusual transcription factor enrichment profile described in the bamFingerprint graph above.

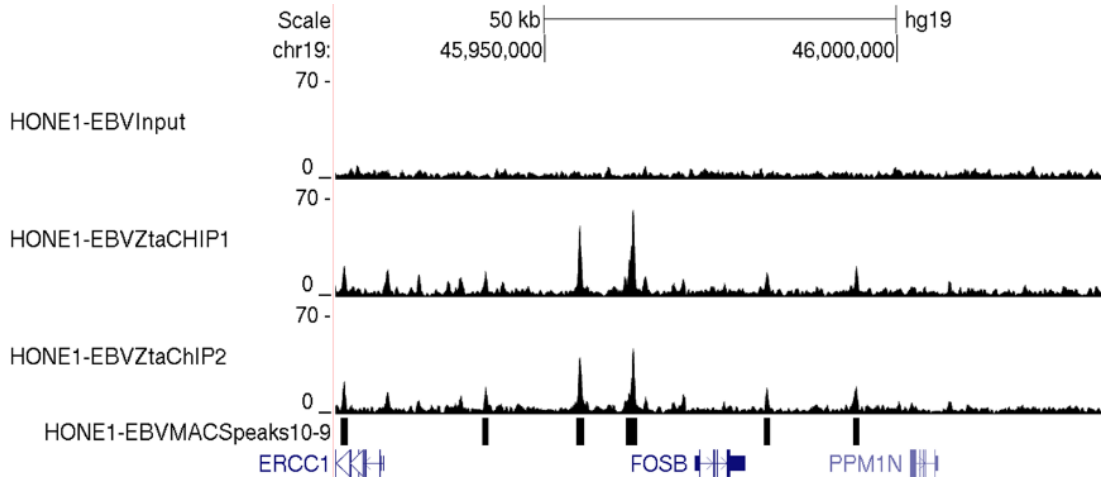


Figure 34: Example of HONE1-EBV Zta ChIP sequencing reads around the FOSB gene. Peaks were visualised in the form of histograms and MACS peaks with the help of the UCSC genome browser. The height of the graph represents the amount of ChIP DNA fragments that have been uniquely mapped to the particular site of the human genome. Cells were induced with 10μM SAHA for 48 hours. Gene locations are indicated below. FOSB is a known Zta binding site in Akata B cells.

To determine the statistical significance of these enriched binding sites, model-based analysis of ChIP Sequencing (MACS) peak calling was performing using the deepTools Galaxy platform [116]. The number of significant peaks at two different p-values is summarized in Table 5 below, together with the number of

peaks common to both ChIPs (intersected peaks). Peaks were detected using the MACS-1.4.2 program of the deepTools Galaxy platform with parameters as specified in the Methods section. The process of intersecting the overlapping pieces of intervals in both ChIPs helps to narrow down the actual binding region of Zta to the human genome more effectively. Further analysis was carried out based on MACS peaks at a p-value of 10^{-9}.

p-value	ChIP and Input Control	Number of peaks called	Intersected peaks
10^{-7}	CHIP1	18 280	12 378
	ChIP2	13 498	
10^{-9}	ChIP1	14 669	9 913
	ChIP2	10 663	

Table 5: Number of MACS peaks with p-values of 10^{-7} and 10^{-9} in HONE1-EBV Zta ChIP duplicates, and number of peaks present in both ChIPs (intersected peaks). Cells were induced with 10μM SAHA for 48 hours.

Figure 34 shows the histogram of Zta binding sites around FOSB together with MACS peaks at a p-value of 10^{-9}, indicating which of the regions can be considered statistically significant Zta binding sites (HONE1-EBV MACS peaks 10^{-9}). A list of MACS peaks has not been included in the appendix of this thesis due to its unfeasible length. However, the obtained Zta ChIP sequencing data will be publicly available on the Gene Expression Omnibus (GEO) genomics data repository from 01.01.2019 (GSE83354 - Binding site identification for the EBV transcription factor Zta in NPC).

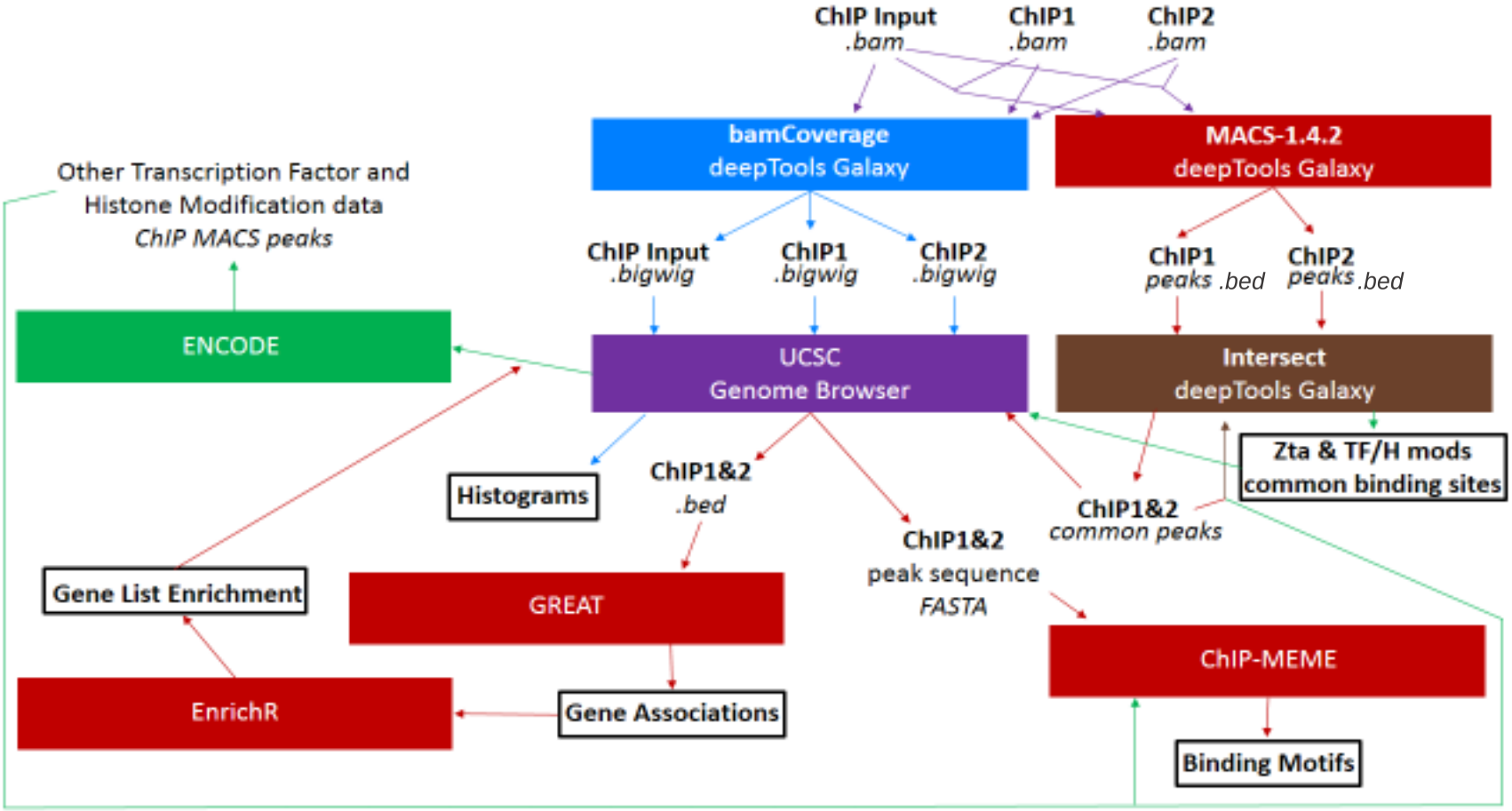


Figure 35: Flow diagram summarising the HONE1-EBV ChIP data analysis. *.bam* files were received from Eurofins Genomics, with sequencing reads mapped to the human genome. Programs, tools and file types used for subsequent data analysis steps are shown.

Figure 35 gives an overview of the data analysis steps, and subsequent results, of the HONE1-EBV Zta ChIP data analysis. Each step will be outlined further in the text below.

For the purpose of creating the histograms shown in Figure 34, binary *.bam* files were binned and converted to *.bigwig* files by normalizing to reads per kilobase per million reads, using the bamCoverage tool of the deepTools Galaxy platform, which reduces the file size so that the data can be loaded and visualised in the University of California, Santa Cruz (UCSC) genome browser.

4.3 Zta binding and host gene regulation

A ChIP qPCR experiment was conducted to confirm the binding of Zta to specific sites on the human genome, which were discovered through ChIP sequencing analysis and are known Zta binding sites in Akata cells. Figure 36 demonstrates the binding of Zta to SCIMP and BCL2A1 promoter regions of the HONE1-EBV cell genome (published in [111]), thus confirming the validity of MACS peak results. It appears that the Zta antibody binds more to non-specific regions than the IgG control antibody, for which no binding could be shown, despite this, enrichment can be seen at peak binding sites in comparison to flanking regions.

To investigate whether genes are regulated as a result of Zta binding peak gene association in HONE1-EBV cells, the level of specific RNAs was quantified with the help of reverse-transcribed cDNA qPCR analysis. These particular genes have been shown to be upregulated during the EBV lytic cycle in RNA sequencing experiments [111] (data for SLC6A7 not published). As shown in Figure 37, the

binding of Zta at these sites is associated with the up-regulation of gene transcription in epithelial cells.

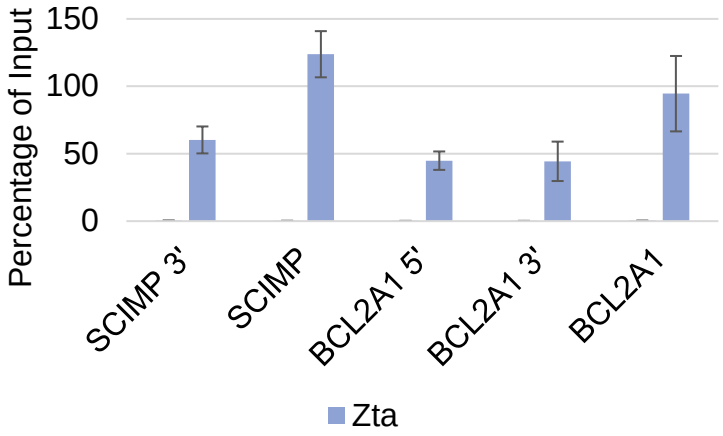


Figure 36: ChIP qPCR analysis showing the amount of HONE1-EBV DNA bound by Zta when induced with 10μM SAHA for 48 hours. Chromatin was immunoprecipitated with an antibody specific to Zta or a negative control IgG antibody. ChIP sequencing data analysis has revealed MACS peaks in the promoter region of the SCIMP and BCL2A1 genes of the human genome. The amount of Zta ChIP DNA in this region, and in a flanking negative control region, were quantified. Data represented as percentage of input. Line indicates background. Values are represented as mean ± SD. Increased binding to SCIMP and BCL2A1 in comparison to the flanking regions is significant ($p<0.05$).

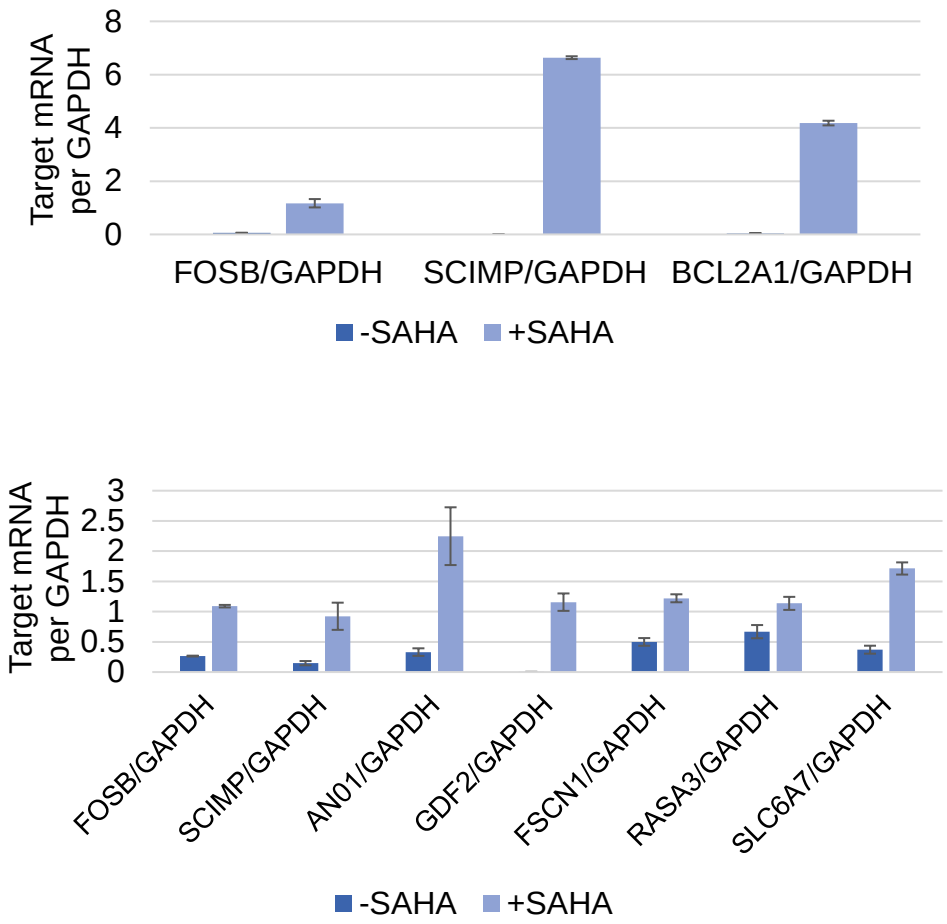


Figure 37: qPCR result showing the amounts of RNA relative to GAPDH levels in HONE1-EBV cells. Lytic cycle was induced (+SAHA) or unindicted (-SAHA) with 10μM SAHA for 48 hours. The increased gene expression in treated cells in comparison to untreated cells is significant for all of these genes except RASA3 ($p<0.05$). Values are represented as mean ± SD.

To investigate whether this binding is specific to HONE1-EBV cells, Zta ChIP qPCR was also performed in EBV positive HEK 293 Zta knock out (ZKO) epithelial kidney cells containing EBV induced into lytic cycle. Results can be seen in Figure 38 below and demonstrate that the level of enrichment of the

SCIMP promoter region is less significant in this cell line, although BCL2A1 is similar to HONE1-EBV.

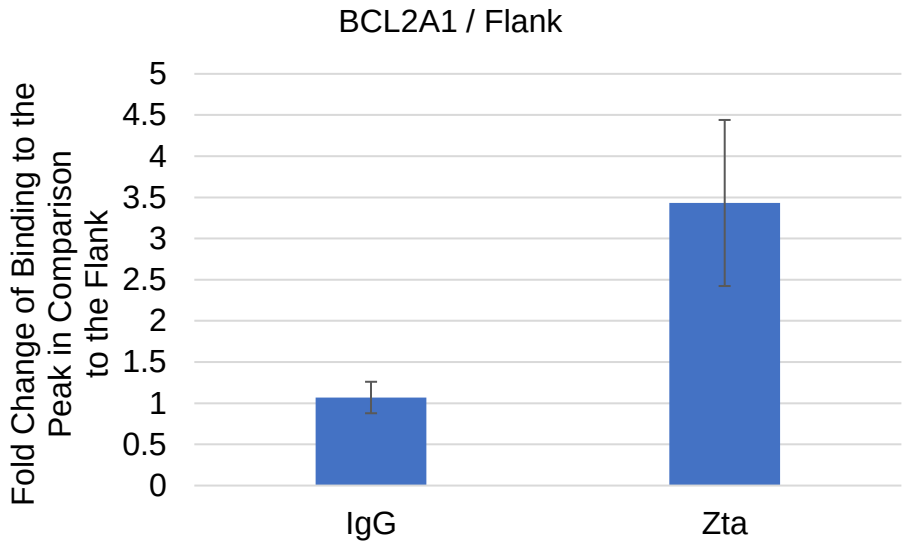


Figure 38: ChIP qPCR analysis showing the relative amount of HEK293 ZKO chromatin bound by Zta when induced into lytic cycle through transfection with Zta. Chromatin was immunoprecipitated with an antibody specific to Zta or a negative control IgG antibody. ChIP sequencing data analysis has revealed MACS peaks in the promoter region of the BCL2A1 gene of the HONE1-EBV genome. The amount of ChIP DNA in this region, and in a flanking negative control region, were quantified. The fold change of binding to these sites is represented in the graph. The enhanced binding of Zta to the peak in comparison to the flanking region is significant ($p<0.005$). Values are represented as mean ± SD.

4.4 Investigation of the mechanism of action by which Zta controls epithelial cells

In order to investigate the mechanism of action through which EBV may be controlling the epithelial host cell, via the viral transcription factor Zta, the binding motif of the protein was investigated. For this purpose we utilised the Multiple Expectation Maximization for Motif Elicitation (MEME) -ChIP binding analysis platform [118] to seek common motifs in FASTA files, obtained from the UCSC Table browser (Figure 35), of all HONE1-EBV Zta ChIP MACS peak sequences. As can be seen in Figure 39, the most common binding motif identified is a previously characterised Zta binding site in B cells which strongly resembles the non CpG Zta response element (ZRE) [79].

A CA repeat element was also identified as a Zta binding motif with a significant p-value of 1.4×10^{-382}. This motif has also been previously identified during similar experiments in a B cell background (data not published).

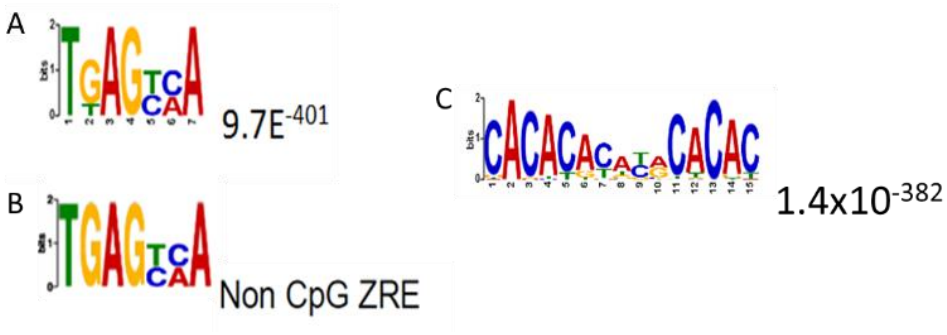


Figure 39: MEME-ChIP analysis of Zta binding motifs. A: Most significant Zta binding motif identified in a Zta ChIP sequencing experiment using HONE1-EBV cells induced with 10μM SAHA for 48 hours. B: Non CpG ZRE previously identified in a Zta ChIP sequencing experiment using EBV positive B cells in lytic cycle. C: A CA repeat element was also identified as a Zta binding motif with a significant p-value of 1.4 x 10^{-382}. This motif has also been previously identified during similar experiments in a B cell background.

To investigate the mechanism of action of Zta on the epithelial cell genome further, we cross-referenced the genes associated with HONE1-EBV Zta ChIP MACS peak obtained from GREAT, with sites of histone modifications and human transcription factor ChIP data in a range of different cell types using the gene list enrichment analysis tool EnrichR [119] which utilises the Encyclopaedia of DNA Elements (ENCODE) database. The list of genes associated with Zta binding peaks in HONE1-EBV cells is too large to be included in this thesis. However, as mentioned previously, the Zta ChIP sequencing data will be publicly available on GEO from 01.01.2019.

The analysis showed a co-association of Zta binding sites with H3K27me3, most significantly in other epithelial cell types (Table 6). H3K27me3 is associated with

inactive gene promoters and shutting down the transcription process. Remarkably, H3K27me3 has only one methyltransferase, namely EZH2, which is part of the PRC2 complex [135].

Histone Modification_Cell Line_Genome Build	Overlap	P-value
H3K27me3_BJ_hg19	107/2090	$2.32x10^{-28}$
H3K27me3_bronchial epithelial cell_hg19	103/2082	$4.09x10^{-26}$
H3K27me3_fibroblast of lung_hg19	125/3052	$7.48x10^{-25}$
H3K27me3_kidney epithelial cell_hg19	101/2217	$5.46x10^{-23}$
H3K27me3_GM12878_hg19	126/3221	$2.34x10^{-23}$
H3K27me3_endothelial cell of umbilical vein_hg19	127/3393	$6.17x10^{-22}$
H3K27me3_mammary epithelial cell_hg19	107/2586	$2.51x10^{-21}$
H3K27me3_myocyte_mm9	67/1183	$9.73x10^{-20}$
H3K27me3_cardiac mesoderm_hg19	126/3485	$1.67x10^{-20}$

Table 6: Table showing the co-associations of Zta ChIP sequencing data of HONE1-EBV cells. Lytic cycle was induced with 10μM SAHA for 48 hours. Histone modification data derived from the ENCODE database. Gene list enrichment analysis performed by the EnrichR tool.

EnrichR analysis also showed a significant co-association between Zta and a range of host cell “transcription factors” in B cells and other cell types (Table 7). While these proteins are classified in the transcription factor data base, most of the proteins listed here do not bind DNA directly.

Notably, EZH2 and SUZ12 are part of the same PRC2 complex which is involved in H3K27 trimethylation.

Transcription Factor_Cell Line_Genome Build	Overlap	P-value
BCL3_GM12878_hg19	75/503	$1.85x10^{-44}$
EZH2_B cell_hg19	88/2000	$3.55x10^{-16}$
SUZ12_H1-hESC_hg19	106/2827	$6.45x10^{-15}$
SUZ12_NT2-D1_hg19	35/563	$1.49x10^{-10}$
EZH2_skeletal muscle myoblast_hg19	71/2000	$4.36x10^{-09}$

Table 7: Table showing the overlap of Zta ChIP sequencing data of HONE1-EBV cells. Lytic cycle was induced with 10μM SAHA for 48 hours, and human transcription factor ChIPs in various cell types. Human transcription factor ChIP data derived from the ENCODE database. Gene list enrichment analysis performed by the EnrichR tool.

Using the Genomic Regions Enrichment of Annotations Tool (GREAT) [120], we found that most of the Zta binding sites on the epithelial cell genome are more than 2kb away from transcription start sites (Figure 40). To generate the graph, we analysed the data repeatedly, changing the parameter of the two nearest transcription start sites within a certain kilo base distance from the peak. The resulting number of peaks was plotted. This showed that Zta may act through distal regulatory elements to control host cell gene as has been shown to be the case in B cells [105].

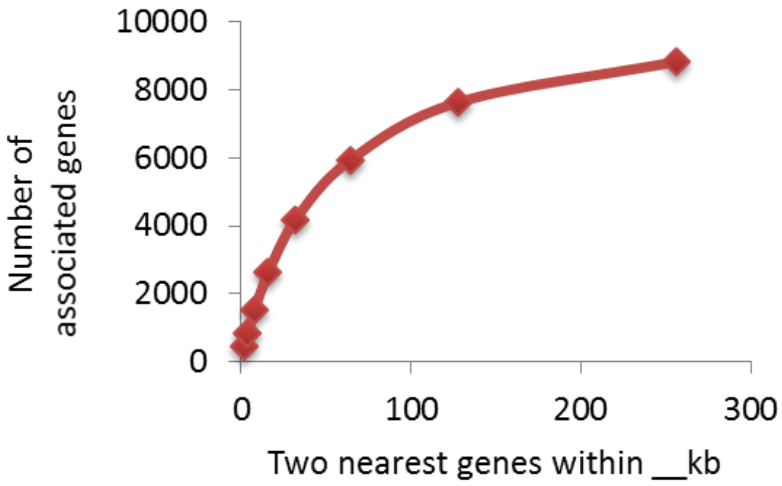


Figure 40: Graph showing the number of genes, within a certain distance (in kb), associated with Zta ChIP sequencing MACS peaks. Lytic cycle was induced in HONE1-EBV cells with 10μM SAHA for 48 hours. To generate the grap the data was analyzed repeatedly, changing the parameter of the two nearest transcription start sites within a certain kilo base distance from the peak. The resulting number of peaks was plotted. Graph shows that most of the Zta binding sites on the epithelial cell genome are more than 2kb away from transcription start sites.

With the help of GREAT we were able to identify pathways which might be associated with Zta host cell binding sites. For this purpose, GREAT utilises different databases containing the Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system [136] and the Gene Ontology (GO) enrichment analysis tool [137]. Identified pathways include the notch signalling pathway (p-value: 1.84×10^{-8}), negative regulation of receptor activity (p-value: 1.06×10^{-7}), positive regulation of steroid biosynthesis (p-value: 1.15×10^{-7}) and steroid metabolic processes (p-value: 3.55×10^{-6}). Considering binding

sites within 2kb of the transcription start site only, which is regarded as the promoter region, we discovered associations with the negative regulation of receptor activity (p-value: 1.06×10^{-7}), positive regulation of steroid biosynthesis (p-value: 1.15×10^{-7}) and steroid metabolic processes (p-value: 3.44×10^{-6}). The list of genes associated with Zta peaks within 2kb of their promoter region has been included in Appendix A of this thesis (Supplementary Figure 1).

Biological Process	p-value
oxygen transport	1.01×10^{-5}
gas transport	2.86×10^{-5}
behavior	6.66×10^{-5}
neuron differentiation	2.17×10^{-4}
inner ear morphogenesis	2.19×10^{-4}

Cellular Component	p-value
hemoglobin complex	3.58×10^{-5}
extracellular matrix part	2.88×10^{-5}
neuronal postsynaptic density	1.63×10^{-3}
receptor complex	3.64×10^{-3}
neuron spine	3.09×10^{-3}

Molecular Function	p-value
oxygen transporter activity	5.35×10^{-6}
oxygen binding	5.29×10^{-4}
heme binding	8.90×10^{-4}
iron ion binding	1.65×10^{-3}
tetrapyrrole binding	1.44×10^{-3}

Table 8: Top five most significant gene ontologies linked to genes associated with Zta ChIP sequencing MACS peaks in HONE1-EBV cell. Lytic cycle was induced with 10μM SAHA for 48 hours. All gene ontologies (GOs) are at a low level of significance. A recurring theme in GOs linked to genes associated with Zta MACS peaks in HONE1-EBV cells seems to be oxygen transport related.

EnrichR was used to further explore the function of Zta controlling the host cell. This tool enabled us to investigate gene ontologies (GO) which may be linked to genes associated with Zta MACS peaks in the context of biological processes,

cellular components and molecular functions (Table 8). All of the obtained results are at a very low level of significance. A recurring theme in GOs linked to genes associated with Zta MACS peaks in HONE1-EBV cells seems to be oxygen transport related.

4.5 Comparison of Zta peaks

To investigate cell type specific properties of Zta, the Zta ChIP sequencing results obtained in HONE1-EBV cells were compared with Zta CHIP sequencing results in Akata cells. MACS peaks for Zta binding to the human genome, which are common to HONE1-EBV and Akata cells, will be among the data published on GEO (as stated in section 4.2).

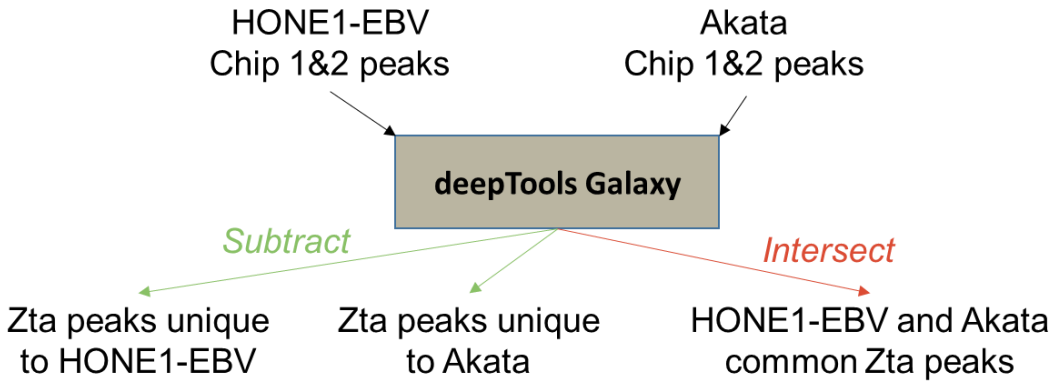


Figure 41: Comparison of Zta peaks in Akata and HONE1-EBV cells. The Zta ChIP sequencing results obtained in HONE1-EBV cells were compared with Zta CHIP sequencing results in Akata cells. This Figure describes the analysis step undertaken during this process.

Figure 41, 43 and 46 highlight specific processes shown in Figure 35. Figure 41 focuses on how the Zta peak data of binding to the host genome of epithelial and

B cells were processed to find binding sites which are common to both cells types or unique to either.

Figure 42 illustrates the overlap of Zta MACS peaks in HONE1-EBV cells with Zta MACS peaks in the B cell line Akata. This demonstrates that approximately 2/3 of Zta MACS peaks in HONE1-EBV cells (NPC) are unique to this cell type in comparison to Zta peaks in Akata B cells, and 3704/5020 Zta peaks in Akata cells are also identified in HONE1-EBV.

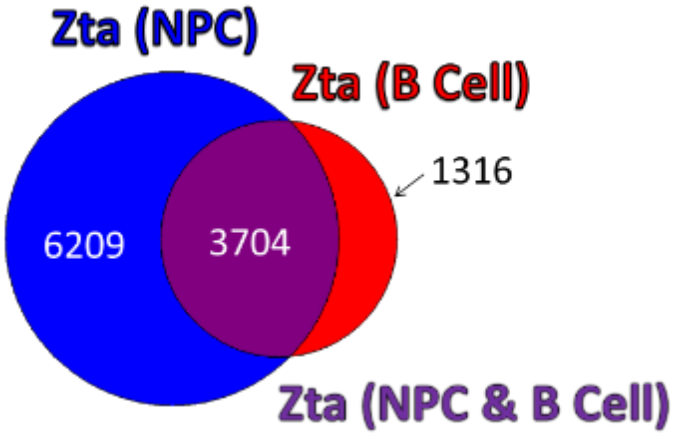


Figure 42: Venn Diagram showing the overlapping pieces of Zta MACS peaks between HONE1-EBV nasopharyngeal carcinoma cells (NPC) and Akata B cells. Zta MACS peaks in HONE1-EBV cells at a p-value of 10^{-9} and Zta MACS peaks in the B cell line Akata at a p-value of 10^{-7}. This demonstrates that approximately 2/3 of Zta MACS peaks in HONE1-EBV nasopharyngeal carcinoma cells (NPC) are unique to this cell type in comparison to Zta peaks in Akata B cells.

Figure 43 summarises how specific Zta were analysed further to obtain binding motifs, gene associations and gene list enrichments described in the following section.

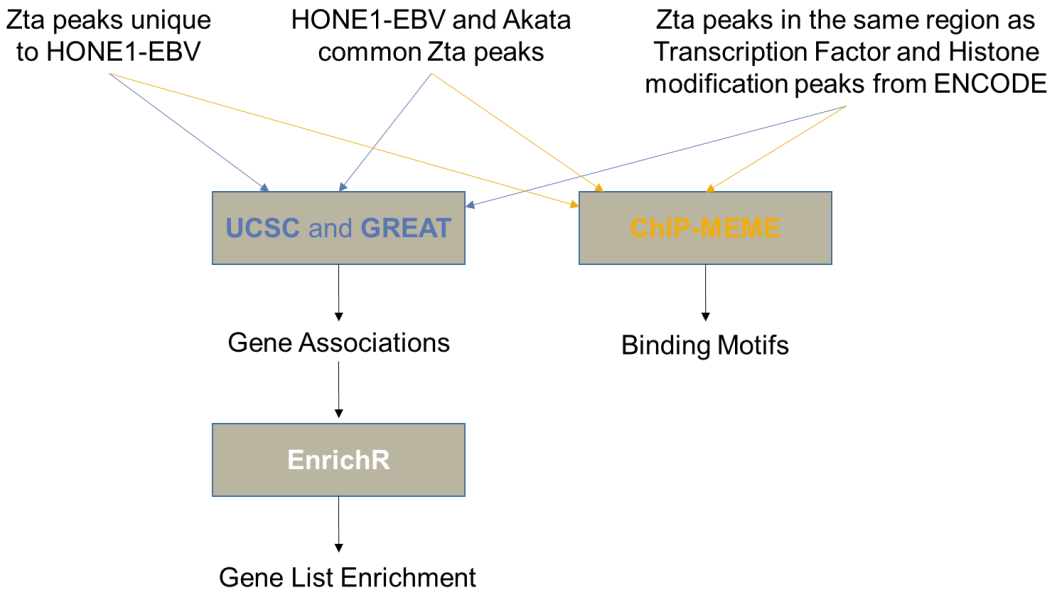


Figure 43: Analysis of Zta peaks in Akata and HONE1-EBV cells. Results obtained from further analysis of Zta peak files are indicated with the analysis tool used stated in boxes.

Oxygen transport related GO analysis results can also be observed linked to Zta MACS peaks present in HONE1-EBV cells only, not present in Akata B cells. However, looking at genes associated with Zta MACS peaks present in both cell types does not yield oxygen transport related GOs, indicating that this analysis might point to a particular Zta function specific to epithelial cells. The list of genes associated with Zta peaks in HONE1-EBV and Akata cells is too large to be included in this thesis.

Focusing in more closely on Zta MACS peaks which NPC and B cells have in common, ChIP-MEME analysis was performed on the intersecting peak FASTA files. This analysis revealed that the most significant Zta binding motif associated with MACS peaks present in NPC and B cells is a CA repeat element ($7.1x10^{-169}$). The Non-CpG ZRE described earlier, as well as a CpG ZRE were identified (Figure 44). The p-values for these motifs are substantially less significant than those obtained for motifs found in HONE1-EBV cells (Figure 39) because the number of peaks put in to this analysis is much lower.

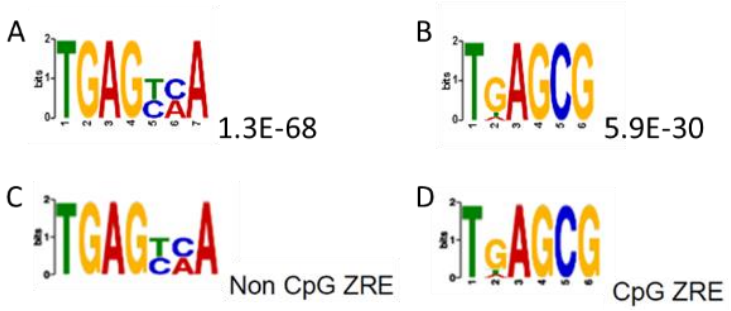


Figure 44: MEME-ChIP analysis of Zta binding motifs. A: Significant Zta binding motif identified in Zta ChIP sequencing MACS peaks in HONE1-EBV cells, induced with 10μM SAHA for 48 hours, intersected with Akata B cell peaks induced with 0.125% IgG for 24 hours. **B**: Significant Zta binding motif identified in Zta ChIP sequencing MACS peaks in HONE1-EBV cells, induced with 10μM SAHA for 48 hours, intersected with Akata B cell peaks induced with 0.125% IgG for 24 hours. **C**: Non CpG ZRE previously identified in a Zta ChIP sequencing experiment using EBV positive B cells in lytic cycle. **D**: CpG ZRE previously identified in a Zta ChIP sequencing experiment using EBV positive B cells in lytic cycle.

Using GREAT, we found that most of the Zta binding sites present on the NPC cell as well as the B cell genome are more than 2kb away from a transcription start sites. The number of genes, within a certain distance (in kilo bases), associated with Zta ChIP sequencing MACS peaks in epithelial and B cells resemble that shown in Figure 40. GREAT analysis of the associated genes shows that pathways which might be involved in the mechanism of Zta action include cell regulation and differentiation gene sets like cell differentiation in the spinal cord ($3.68x10^{-19}$), positive regulation of protein serine/threonine kinase activity ($2.32x10^{-18}$) and negative regulation of osteoblast differentiation ($7.38x10^{-18}$). A list of genes, associated with Zta peaks common to HONE1-EBV and Akata cells, within 2kb of the transcription start site, is included in Appendix A of this thesis (Supplementary Figure 2).

Genes associated with HONE1-EBV Zta ChIP MACS peaks shared with B cells were cross-referenced with sites of histone modifications and human transcription factor ChIP data from ENCODE, in a range of different cell types, using EnrichR (Figure 43) [119]. The histone modifications analysis showed a most significant association of Zta binding sites with H3K4me1, which has a higher p-value than the most relevant histone modification in the complete set of Zta peaks in NPC cells. H3K4me1 is characterised by post-transcriptional modifications and associated with active enhancers [138].

A less significant association with H3K27me3 is present, when compared to results from the complete set of Zta peaks in NPC cells. This data could suggest

that Zta fulfils a role specific to NPC cells, which differs from its mechanism of action in B cells, making H3K27me3 more relevant in this cell type.

In addition, a significant co-association with the same top three transcription factors as in the previous analysis was revealed (Table 9). The overlap with BCL3 seems to be even more significant with Zta peaks common to NPC and B cells (3.18×10^{-52}), this could suggest a cell type independent role of this particular transcription factor binding site to the lytic cycle of EBV. Interestingly, BCL3 has been shown to interact with CREB during its role as transcription repressor and the CREB binding protein (CBP) has also gained in relative significance during this particular co-association analysis.

Transcription Factor_Cell Line_Genome Build	Overlap	P-value
BCL3_GM12878_hg19	54/503	3.18×10^{-52}
EZH2_B cell_hg19	40/2000	1.18×10^{-12}
SUZ12_H1-hESC_hg19	45/2827	6.68×10^{-11}
CREBBP_K562_hg19	30/2000	7.38×10^{-07}
REST_Panc1_hg19	12/688	5.78×10^{-4}

Table 9: Table showing co-association of Zta ChIP sequencing data in HONE1-EBV cells intersected with Akata cells, and human transcription factor ChIPs in various cell types. Lytic cycle in HONE1-EBV was induced with 10μM SAHA for 48 hours. Lytic cycle in Akata cells was induced with 0.125% IgG for 24 hours. Human transcription factor ChIP data derived from the ENCODE database. Gene list enrichment analysis performed by the EnrichR tool.

4.6 A closer look at Zta peaks at host transcription factor promoter regions

Based on the EnrichR analysis, we investigated the link between the viral transcription factor (TF) Zta and the human TFs BCL3, EZH2 and SUZ12 further. The MACS peak binding profile of Zta in the region of genes encoding these TFs is shown in Figure 45. The presence of binding sites near the genes suggests the possibility of a feedback mechanism between Zta and these host cell TFs in all three cases for epithelial cells, and for EZH2 also in B cells.

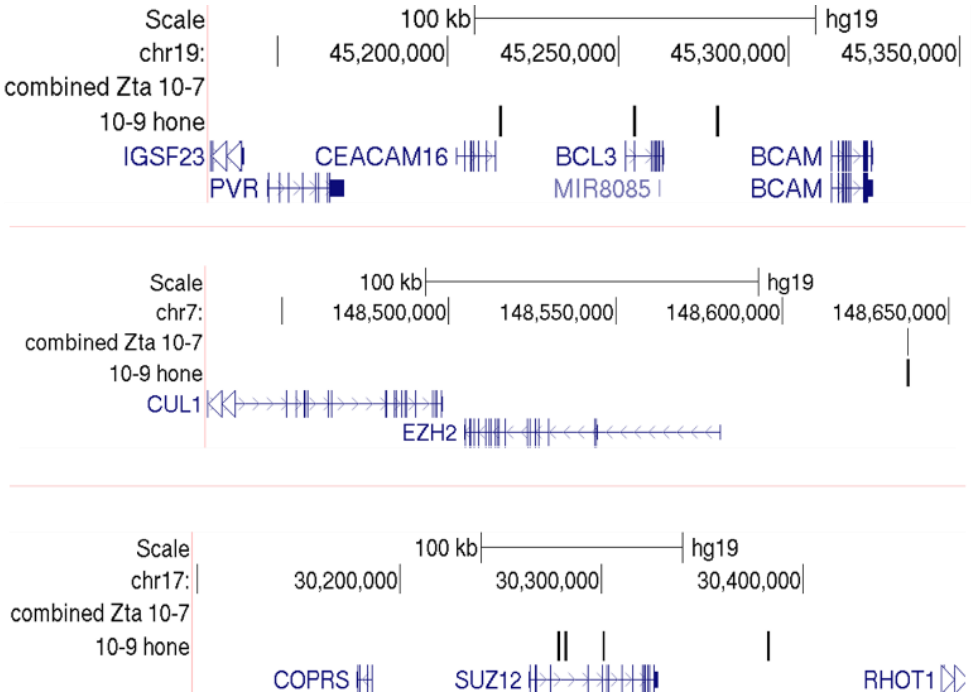


Figure 45: Zta MACS peak binding sites in the region of significant human TF genes. HONE1-EBV cells were induced with 10μM SAHA for 48 hours (10-9 hone) and Akata cells induced with 0.125% IgG for 24 hours (combined Zta 10-7).

Having found significant co-associations of Zta binding peaks with a range of other protein and histone modification peaks, the possibility of an absolute

overlap of ChIP sequencing peaks was investigated. Figure 46 further expands on the analysis steps by showing how transcription factor and histone modification data from ENCODE were included to extend the investigation of the mechanism of action of Zta on the host cell genome.

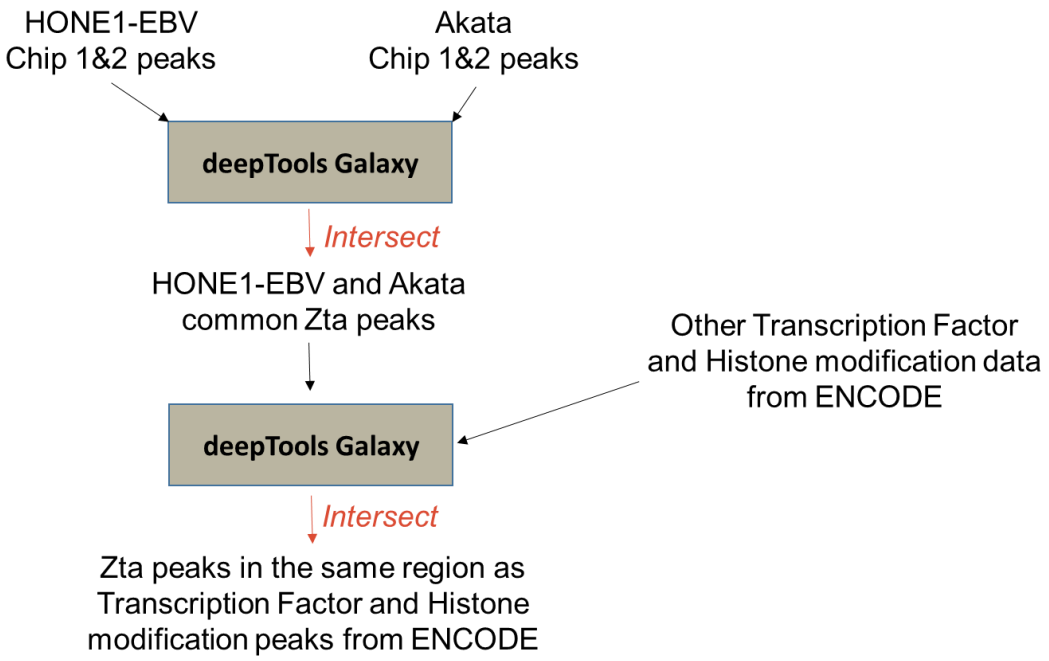


Figure 46: Comparison of Zta peaks in Akata and HONE1-EBV cells with other transcription factor and histone modification ChIP Seq data. The Zta ChIP sequencing results obtained in HONE1-EBV cells were compared with Zta CHIP sequencing results in Akata cells. The resulting common peaks were compared to other published ChIP sequencing peaks for transcription factors and histone modification. This Figure describes the analysis step undertaken during this process.

Zta peaks found in epithelial cells, Akata B cells, and both cell types were intersected with the BCL3, SUZ12 and EZH2 binding sites, as well as H3K27me3 sites. The results of these intersections are summarised in Table 10.

Site	B & NPC cell peaks	NPC cell peaks	B cells peaks
BCL3	64%	40%	53%
EZH2	0.19%	0.54%	0.30%
H3K27me3	44%	45%	44%
SUZ12*	61%	54%	62%

* Data not available in GM12878 cells, values relate to H1hESC cells

Table 10: Percentage of Zta peaks in HONE1-EBV NPC cells which overlap with indicated cellular factors or sites of H3K27me3. Lytic cycle in HONE1-EBV cells was induced with 10μM SAHA for 48 hours. Lytic cycle in or Akata B cells was induced with 0.125% IgG for 24 hours.

A high number of sites which overlap with Zta and a human transcription factor can be seen for BCL3. SUZ12 has similarly high overlap with Zta binding sites in NPC and B cells, as does H3K27me3. Though EZH2 has only a small amount of overlap with Zta peaks.

The actual numbers of BCL3 peaks in GM12878 cells which overlap with Zta peaks in HONE1-EBV, Akata cells, and Zta peaks common to both is pictured in the Venn Diagrams in Figure 47.

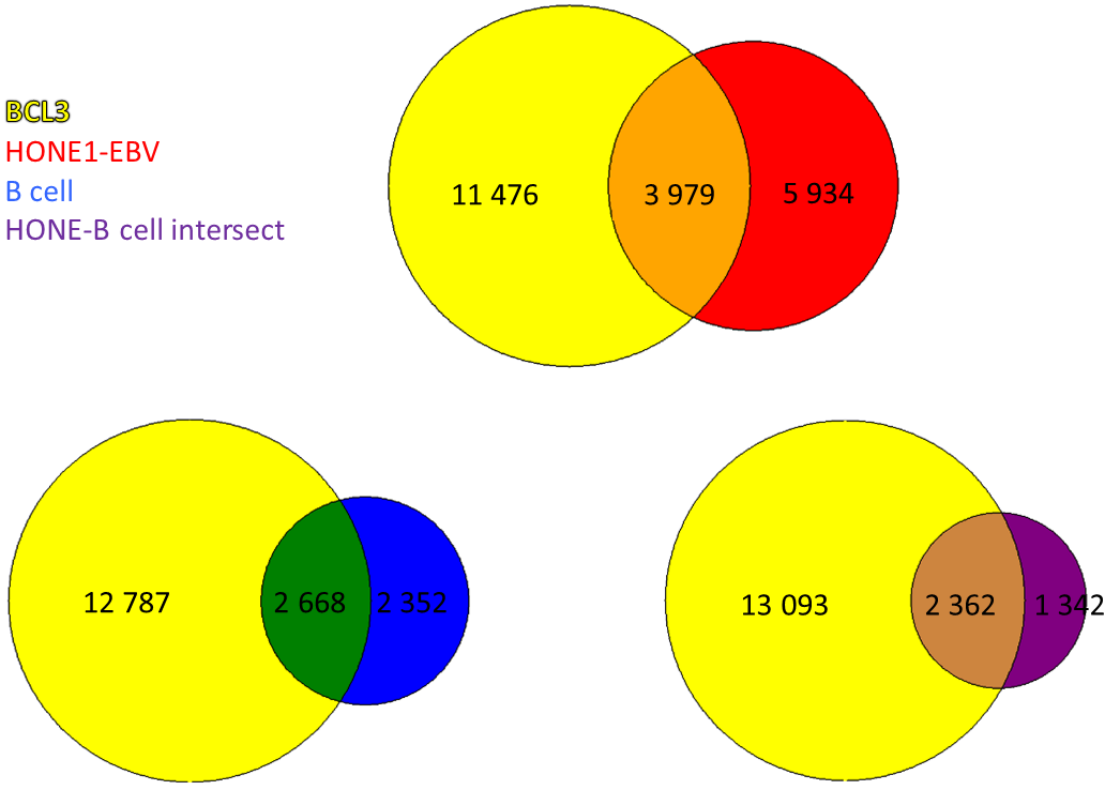


Figure 47: BCL3 peaks intersected with Zta peak. BCL3 ChIP sequencing peaks in GM12878 cells were intersected with Zta peaks in HONE1-EBV, Akata and common binding peaks.

In order to investigate the link between BCL3 and Zta further, ChIP-MEME analysis was carried out to identify the BCL3 binding motif. For this purpose publicly available ChIP sequencing data was used in the GM12878 and K562 cell backgrounds. K562 are EBV negative erythroleukaemia cells while GM12878 is an EBV positive lymphoblastoid cell line. Results of this analysis can be seen in Figure 48.

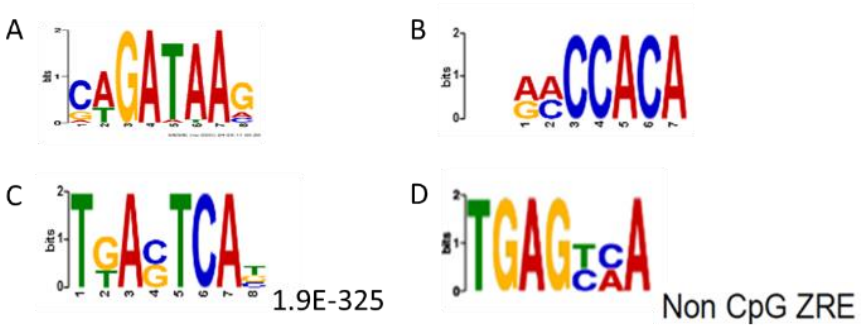


Figure 48: MEME-ChIP analysis of indirect BCL3 binding motifs. A: BCL3 binding motif identified in BCL3 ChIP sequencing MACS peaks in EBV negative K562 cells in the factorbook repository. B: CA rich motif identified in BCL3 ChIP sequencing MACS peaks in EBV positive GM12878 cells by ChIP-MEME analysis. C: Most significant BCL3 binding motif identified in BCL3 ChIP sequencing MACS peaks in EBV positive GM12878 cells by ChIP-MEME analysis. D: Non CpG ZRE previously identified in a Zta ChIP sequencing experiment using EBV positive B cells in lytic cycle.

The BCL3 binding motif in GM12878 cells, an EBV positive lymphoblastoid cell line, is similar to the Non CpG ZRE, further supporting the proposal that BCL3 could play an important role in the regulation of the host cell by EBV. A CA rich motif was also discovered in the GM12878 BCL3 binding region. The Non CpG ZRE cannot be identified in the EBV negative K562 cell line.

4.7 Discussion

The mapped ChIP sequence reads initially suggest a relatively poor transcription factor enrichment on the bamFingerprint graph. However, this is due to a clustering of Zta binding peaks which leads to an enrichment pattern more

characteristic of broad histone modification peaks, rather than high, narrow peaks associated with most transcription factors [139]. In addition, MACS peaks were called at a high significance with a p-value of 10^{-9}, which point to ample enrichment of the sequence reads. Peaks were also called at a lower significance with a p-value of 10^{-7}, which yielded more peaks, however, visual inspection of genome alignments confirm that these are lower quality peaks. Further investigations focused on the more significant peaks only. While some relevant peaks might be missed at the higher significance, there is still a large number of more significant peaks to examine which are more likely to be pertinent.

Several HONE1-EBV Zta MACS peaks were confirmed, with the help of ChIP qPCR, at locations which are known Zta binding sites in Akata cells [105]. Enrichment of the same regions can also be observed in a second epithelial cell line, HEK293 ZKO, though the effect was somewhat less substantial.

Motif analysis of Zta binding sites revealed the non CpG ZRE, which was previously shown to be a Zta binding motif in Akata cells [18]. However, the CpG ZRE could not be detected, even at a low significance, for Zta binding in epithelial cells. This could point to a B cell specific function of Zta linked to CpG ZREs.

To investigate if the binding of Zta to the epithelial cell genome has an effect on gene transcription, the RNA levels of associated genes were quantified. This analysis showed a significant up-regulation of several RNA levels in comparison to uninduced cells. This result is an average of the whole cell population, in which only approximately 21% of the cells are actually expressing Zta (Chapter 3).

A closer look at gene associations revealed that the majority of Zta peaks are further than 2kb away from a transcription start sites, suggesting that Zta acts

through distal regulatory elements in epithelial cells, as it does in B cells [105]. The pathways linked to Zta gene associations should be further investigated, with a closer look at the possible regulation of common genes.

The oxygen transport related GOs outlined, together with the positive regulation of steroid biosynthesis and steroid metabolic processes, might point to an increase in cell metabolism. This set of GOs seems specific to EBV infection in epithelial cells, as oxygen transport related results can also be found for Zta peaks unique to epithelial cells, but not for peaks common to both cell types. In addition, these metabolic and biosynthetic pathways are not associated with peaks common to both cell types. However, as is shown, the number of significant peaks in B cells, even at a lower significance of 10^{-7}, is smaller than that in epithelial cells. This is due to a lower sequencing depth in B cells, which may have led to the exclusion of some relevant peaks in this cell type.

Gene association analysis indicates a significance of H3K27me3 linked to the role of Zta in controlling the epithelial host cell. In addition, we also found a significant co-association between Zta, EZH2 and SUZ12, which are part of the PRC2 complex responsible for H3K27 trimethylation [140]. HE327me3 is associated with silenced regions of the genome. The relationship between methylated DNA and H3K27me3 is still poorly defined. In embryonic stem cells, H3K27me3 is located almost exclusively at CpG islands which are generally not associated with DNA methylation. However, in cancer cell lines H3K27me3 is a lot less restricted to CpG islands and there are extensive overlap between H3K27me3 and methylated DNA [141, 142]. Interestingly, it has previously been shown that the viral transcription factor, Zta, is able to bind to methylated DNA [76].

The most significant transcription factor co-association has been found to be with BCL3. For peaks common to both, epithelial and B cells, the repressive chromatin mark H3K27me3 drops in significance. The most significance histone modification for this set of peaks is the active chromatin mark H3K4me1. EZH2 and SUZ12 also drop in significance for common peaks, in comparison to all HONE1-EBV cell peaks, while BCL3 gains in significance. This data could suggest that Zta fulfils a role specific to epithelial cells, which differs from its mechanism of action in B cells, making H3K27me3 more relevant in this cell type. Zta binding sites can be found at the promoter regions of BCL3, EZH2 and SUZ12 in epithelial cells and at EZH2 in B cells. This points to a possible feedback mechanism between Zta and these proteins. Table 8 shows that there is a significant absolute overlap between Zta binding sites on the human genome and the binding sites of BCL3 and SUZ12, as well as H3K27me3 sites. EZH2 shows no significant direct overlap, however, it is possible that the binding regions are very close to each other as opposed to direct overlaps, especially as it forms a complex with SUZ12 .

BCL3 has been shown to associate with the same non CpG ZRE binding motif as Zta, in GM12878 cells. GM12878 cells are infected with EBV, but the virus is reportedly latent and does not express Zta. The EBV negative cell line K562 has not been found to contain this binding motif. Further investigation is needed to determine if there is a link between EBV infection and BCL3 activity. Outlined results suggest that BCL3 could be acting as a repressor during EBV infection, being involved in the establishment of a repressive chromatin environment, which is overturned by Zta during the lytic cycle.

The p-value of the presence of the non-CpG ZRE related motif under BCL3 ChIP sequencing peaks in GM12878 cells is $1.9x10^{-325}$. The p-value of this same Zta binding motif in HONE1-EBV cells is $9.7x10^{-401}$. However, when looking only at peaks which are bound by Zta in HONE1-EBV cells as well as Akata cells, the significance falls to $1.3x10^{-68}$. This 5.9 fold decrease in significance might indicate that the combination of BCL3 and the non-CpG ZRE play a more prominent role in promoting the control of the epithelial host cell by EBV than they do in B cells. This fall in significance is, at least partially, linked to the decrease of the overall number of peaks bound by Zta in both cell types. However, despite this we also found a 2.6 fold increase in the significance of CA repeat regions in sites bound by Zta in both cell types in comparison to all sites bound by Zta in HONE1-EBV cells. The significance of CA repeat regions under BCL3 ChIP sequencing peaks is 16.4 fold lower than that of the Zta binding sites common to NPC and B cells. This could indicate a central role of CA repeat regions to EBV infection, unrelated to BCL3.

Further investigations into the role of BCL3 linked to the early lytic cycle of EBV infection are outlined in the following chapter.

Chapter 5: BCL3

5.1. Introduction

This chapter will focus on the relevance of the B-cell lymphoma 3 protein (BCL3) for lytic cycle reactivation of EBV by Zta. As discussed in the previous chapter, BCL3 binds to 64% of the sites also bound by Zta in epithelial and B cells. In addition, BCL3 binding sites were shown to resemble non CpG ZREs in cells infected with EBV.

BCL3 is located at the chromosomal site 19q13, which frequently translocates into the immunoglobulin alpha-locus in some cases of B cell leukaemia (e.g. chronic lymphocytic leukaemia and some lymphomas) [143, 144]. The BCL3 gene itself consists of 9 exons, spanning 11.5kb. There are nine splice variants of BCL3, seven of which are non-protein coding. The two protein coding, variant BCL3 proteins are 308aa and 454aa in length [145]. BCL3 is a phosphoprotein and has a molecular weight of 32kDa (308aa variant) – 47kDa (454aa variant) [144]. In addition, the BCL3 protein also differs in its phosphorylation states, mainly at the serine-rich carboxyl terminus [144].

Both variants of BCL3 contain seven ankyrin repeats. This structure is characteristic for family members of the Inhibitor of κ-B (IκB) proteins, to which BCL3 has close homology. Both IκB and BCL3 interact with the nuclear factor kappa-light-chain-enhancer of activated B cells (NFκB) via their ankyrin tandem repeat motif. The NFκB protein family is central to inflammation. It is present in the cytoplasm of all cells, where it is generally complexed to IκB. In response to inflammatory signals, IκB is phosphorylated and degraded, which leads to the translocation of NFκB to the nucleus where it binds to DNA and activates genes

leading to inflammatory response [146]. In mammals, the NFκB family consists of p65 (RelA), RelB, c-Rel, p50 (p105, NFKB1), and p52 (p100, NFKB2). The subunits associate with each other to form homo- and heterodimers which control gene transcription. All five family members share a conserved 300aa long amino-terminal Rel homology domain, which is required for interaction with IκB, dimerization, DNA binding and nuclear translocation [147].

Unlike IκB, that inhibits NFκB activity, BCL3 can function as a transcriptional co-activator of NFκB target genes, either directly via its proline-rich amino terminus and proline-/serine-rich carboxyl terminus or indirectly, by removing repressive p50 or p52 homodimers from κB sites [144, 148, 149]. In addition, BCL3 is also able to enhance the binding of p50 and p52 homodimers to certain target sites, leading to transcriptional repression [150, 151]. The mechanisms which lead to either effect of BCL3 are currently still poorly understood, but the effect of BCL3 on transcription activation or repression depends, at least partly, on post-transcriptional modification [144, 147]. BCL3 contributes to the regulation of cell proliferation and is required for the attenuation of pro-inflammatory action of activated macrophages and the survival of activated T cells [144, 152].

BCL3 has also been proposed to have a lymphocyte specific role as an essential negative regulator of Toll-like Receptors (TLRs). It has been proposed that BCL3 may play a role in B cell development by allowing immature B cells to resist apoptosis caused by the DNA damage undergone during immunoglobulin class switch recombination and somatic hypermutation [153].

The varied functions described for BCL3 can partly be attributed to its ability to form complexes with a range of binding partners. Figure 49 shows the BCL3

STRING Interaction Network with 25 known binding partners, including CBP (CREBBP) [15]. Furthermore, BCL3 has been shown to form complexes with the NFκB p52-p52 (NFKB2) homodimer and the NFκB p50-p50 (NFKB1) homodimer [9]. Unlike the other subunits of NFκB, the p50 and p52 proteins lack a transactivation domain. When bound to DNA, these homodimers usually inhibit gene transcription [2]. TNFα, IL1β and phorbolester trigger the formation of BCL3-p50 complexes [10]. Tethering of BCL3 to DNA via the p50 homodimers can allow BCL3 to activate gene transcription [11]. It also interacts with proteins outside of the NFκB family including the transcription factor AP1 (c-Fos and c-jun) whose effect it can increase. Other interactions include Tip60 (histone acetyltransferase), histone deacetylase and CYLD [143, 154, 155].

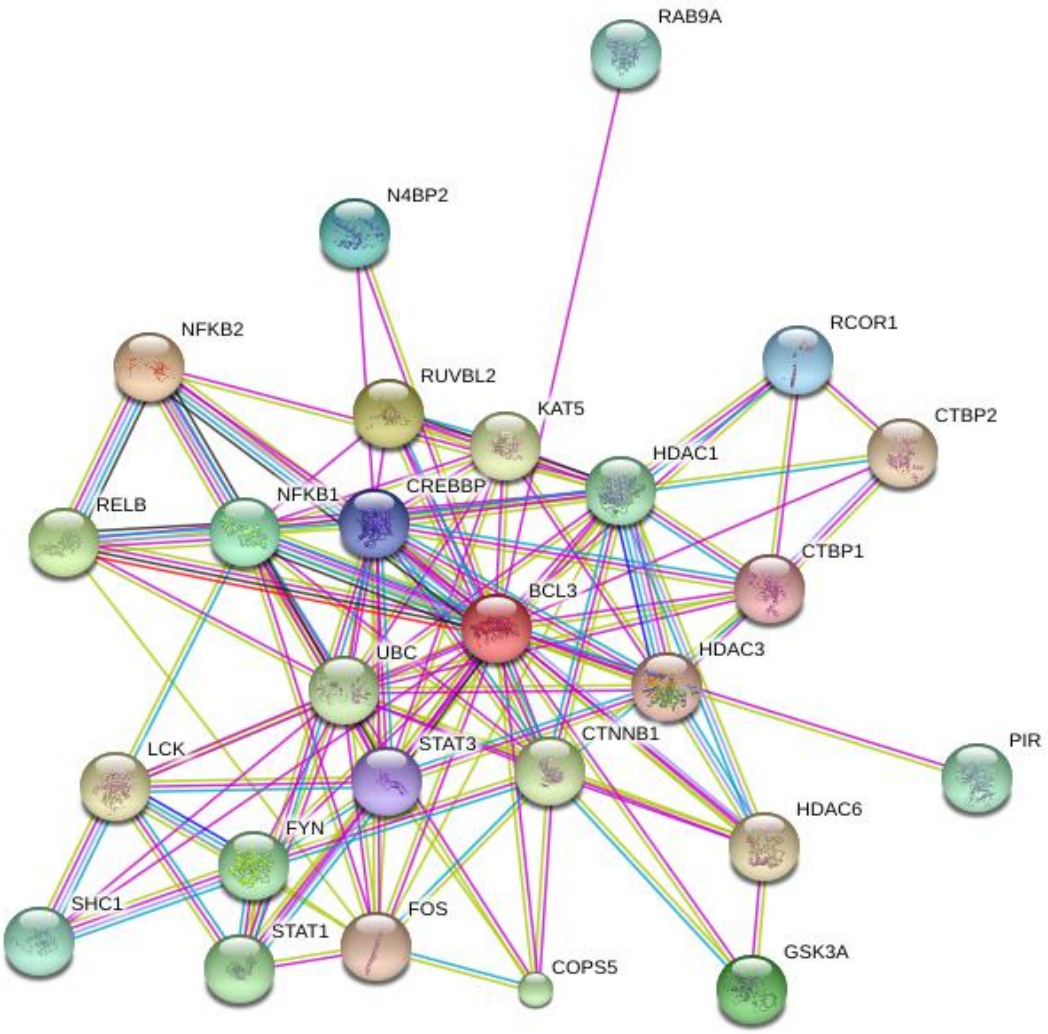


Figure 49: BCL3 STRING Interaction Network. Diagram shows BCL3 with its 25 binding partners [140].

In addition to CBP, BCL3 has also been shown to interact with the transducer of regulated CREB activity 3 (TORC3) during infection with the human T cell leukaemia virus 1 (HTLV-1) [156]. HTLV-1 enters a long period of latency in CD4+ lymphocytes which results in asymptomatic, live-long infection in ~90% of patients. It is associated with adult T cell leukaemia and lymphoma, myelopathy, dermatological conditions, uveitis and subclinical immune suppression which can

result in increased opportunistic infections [157]. The HTLV-1 Tax protein activates the virus' lytic cycle by association with CREB from the viral long terminal repeat which contains multiple cyclic AMP response elements flanked by GC-rich sequences. TORC1, 2 and 3 significantly enhance this Tax-mediated transcription in a manner dependent on a ternary complex composed of Tax, CREB and TORCs. The interaction of BCL3 with TORC3 significantly inhibits the reactivation from latency of HTLV-1. Tax can modulate its own transcription by interacting with CREB, which activated the long terminal repeat [156]. The expression of BCL3 is enhanced by Tax via an NFκB binding site in the second intron of the BCL3 gene [158]. The suppression of transcription by BCL3 is due to the inhibition of TORC3 function, but not inhibition of Tax itself. The amount of CREB recruited to Tax response elements is not altered by increased BCL3 expression [156]. Figure 50 shows a model the activation of transcription by Tax in conjunction with CREB and TORC from the long terminal repeat (LTR) of HTLV-1 (Figure 50 A) and how BCL3 may disrupt the transcription activation activity of Tax by interacting with TORC as part of the Tax-TORC-CREB complex (Figure 50 B), or by inhibiting the binding of TORC and hence the formation of the ternary complex (Figure 50 C).

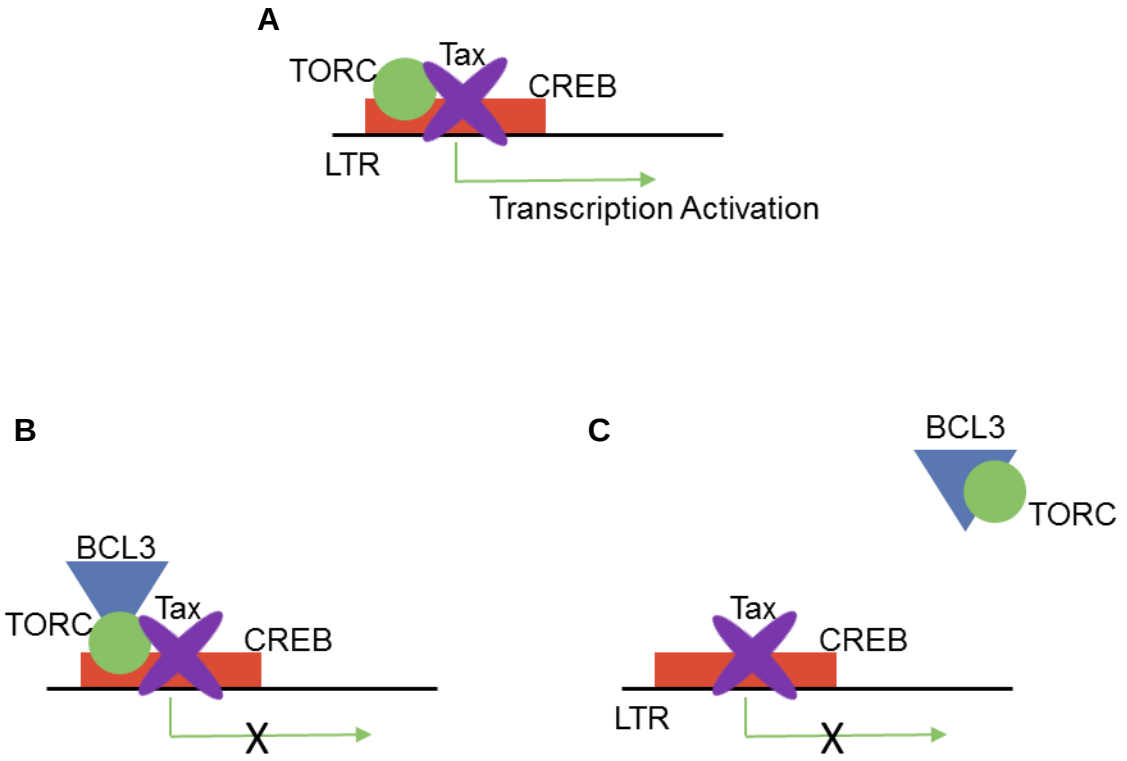


Figure 50: Model of transcription activation by HTLV-1 Tax. A: Activation of transcription by Tax in conjunction with CREB and TORC from the long terminal repeat (LTR) of HTLV-1. The bottom row shows how BCL3 may disrupt the transcription activation activity of Tax by interacting with TORC as part of the Tax-TORC-CREB complex (B), or by inhibiting the binding of TORC and hence the formation of the ternary complex (C).

Evidence indicates that deregulated BCL3 contributes to some types of Hodgkin and non-Hodgkin lymphoma, as well as nasopharyngeal carcinoma and ovarian cancer, where it has been found to be over-expressed without a translocation [143]. The constitutive expression of BCL3 has been shown to suppress DNA

damage-induced p53 activation and consequent p53-induced apoptosis [159]. Additionally BCL3 inhibits apoptosis by blocking Bim activation [160].

The deregulation of BCL3 has also been shown to be associated with a number of other diseases affecting the immune system, including autoimmune type 1 diabetes in mice [161], atopic dermatitis [162] and Crohn's disease [163].

BCL3 abundancy can be regulated at the post-transcriptional level. The translation of BCL3 mRNA transcripts is regulated by Micro-RNAs (miRNAs) or RNA-binding proteins. For example, the translation of BCL3 mRNA is down-regulated by miRNA-125 in human ovarian cancer cells [164].

Surprisingly, transgenic mice, which over-express BCL3 in T and B cells, do not develop a lymphoid neoplasm. However, they do develop a lymphoproliferative disorder. During these experiments, BCL3 was shown to increase the signal-specific survival of follicular B cells as well as the incidence of antibodies to self-antigens. Interestingly, a statistically significant increase in IgA antibodies was also observed [165, 166]. Unfortunately these studies did not include further mouse crosses against oncogenes or knock-outs to test whether BCL3 predisposes to cancer.

BCL3 is expressed in most areas of the human body, but the main defects of BCL3 knock-out mice relate to their immune response. BCL3 deficient mice appear developmentally normal but have an altered spleen and lymph node microarchitecture, including a lack of germinal centre formation. They are unable to form B cell follicles and are not able to form proper follicular dendritic cell networks following antigenic challenge [167]. These knock-out studies in mice

show that the animals have compromised production of antigen-specific B and T cells which makes them vulnerable to certain kinds of pathogens [144].

Epstein–Barr virus latent membrane protein 1 (LMP1) may lead to the deregulation of BCL3 in infected cells. It has been shown that LMP1 expression induces multiple distinct NFκB forms including abundant p50 dimers, p50/p52 dimers and p65. LMP1- carboxy-terminal activating region 1, which is the activating domain of LMP1, also induces formation of a complex containing the p50 dimer and BCL3 on the EGFR promoter. Cells expressing LMP1 have increased levels of phosphorylated STAT3, not only on the BCL3 promoter, but also on the HS3 and HS4 intronic enhancers, resulting in induced BCL3 expression [168]. Indeed, expression of LMP1 is thought to result in p50/BCL3 which is the major form for NFκB detected in Nasopharyngeal Carcinoma [169].

5.2 Expression of BCL3 in a Range of EBV Positive Cell Lines

Western blotting was performed to query whether BCL3 is expressed in HONE1-EBV and Akata cells, which the Zta ChIP sequencing results outlined in the previous chapter are based on. This shows a clear increase in BCL3 protein levels in protein extracts and chromatin of cells containing EBV induced into lytic cycle in comparison to un-induced controls. The detected BCL3 band corresponds to the 47kDa (454aa variant) variant of the BCL3 protein. cDNA qPCR experiments to quantify BCL3 RNA levels did not show an increase in induced samples, suggesting that BCL3 is regulated at the translational level as a result of EBV lytic cycle reactivation or induction.

As the ChIP sequencing results for BCL3 from ENCODE were not based on HONE1-EBV or Akata B cells as the Zta ChIP sequencing results, the expression of BCL3 in these, and other EBV positive cells, was investigated. Figure 51 shows the expression of BCL3, Zta and Actin in HONE1-EBV cell protein extract and chromatin. The Figure shows that BCL3 expression appears to increase when EBV is induced into lytic cycle.

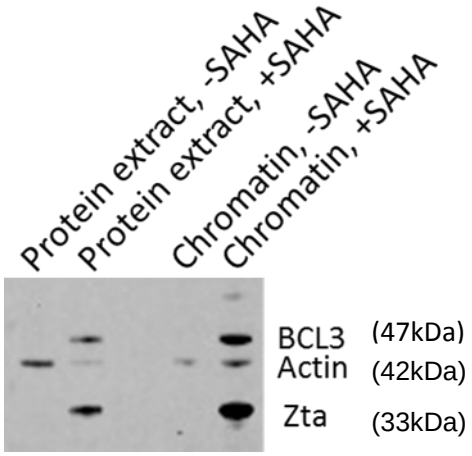


Figure 51: Western Blot showing the expression of Zta, host cell actin and the host transcription co-factor BCL3 in HONE1-EBV cells. Blot shows protein extracts and reverse cross-linked chromatin derived from HONE1-EBV cells un-induced (-SAHA) or induced (+SAHA) into lytic cycle with 10μM SAHA for 48 hours. Blot was performed using the Santa Cruz BCL3 (c-14) antibody.

The corresponding amount of BCL3 RNA in HONE1-EBV cells is shown in Figure 52, along with the RNA level of p50 and p52. There is no significant difference in the amount of BCL3 or p52 RNA, suggesting that any difference in BCL3 protein is likely due to post-transcriptional processing. It is known that BCL3 can be

regulated at the post-transcriptional level by Micro-RNAs (miRNAs) or RNA-binding proteins (eg. miRNA-125) [164].

The amount of p50 RNA increases significantly in HONE1-EBV cells when lytic cycle is induced. This hints at a mechanism of action of lytic cycle reactivation which may involve NFκB, with which BCL3 is known to interact.

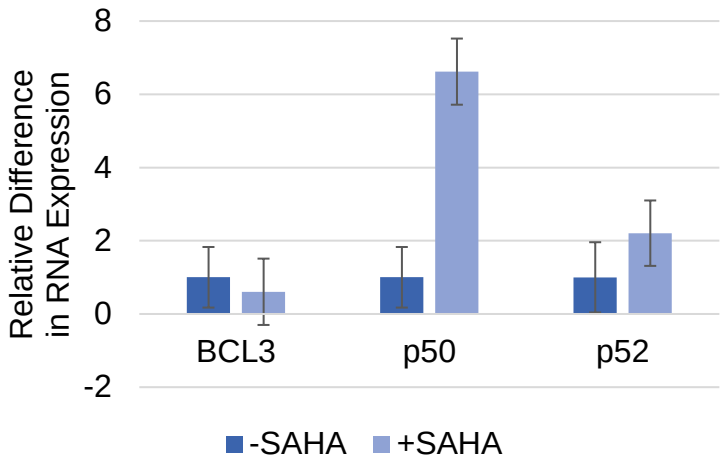


Figure 52: qPCR result showing the fold change of BCL3, p50 and p52 RNA expression in HONE1-EBV cells. Lytic cycle was induced in HONE1-EBV cells with 10μM SAHA for 48 hours (+SAHA) or not induced (-SAHA). Values are represented as mean ± 95% Confidence Interval ($p<0.05$ for p50).

Due to the limitations of the HONE1-EBV cell line, including the fact that it is not fully identified and the drastic decrease in adherent cell numbers when EBV lytic cycle is induced, it was decided to carry out further experiments in alternative cell lines.

Previous Zta ChIP sequencing was also performed in Akata B cells. The last chapter showed a substantial overlap with BCL3 binding sites in this cell type. The expression of BCL3 in Akata cells, as well as HEK293 ZKO cells, can be seen on the western blot in Figure 53. EBV lytic cycle in Akata cells was induced via the activation of a doxycycline inducible promoter, which leads to the expression of Zta, or by crosslinking of the B cell receptor with anti-immunoglobulin [170]. Lytic cycle in

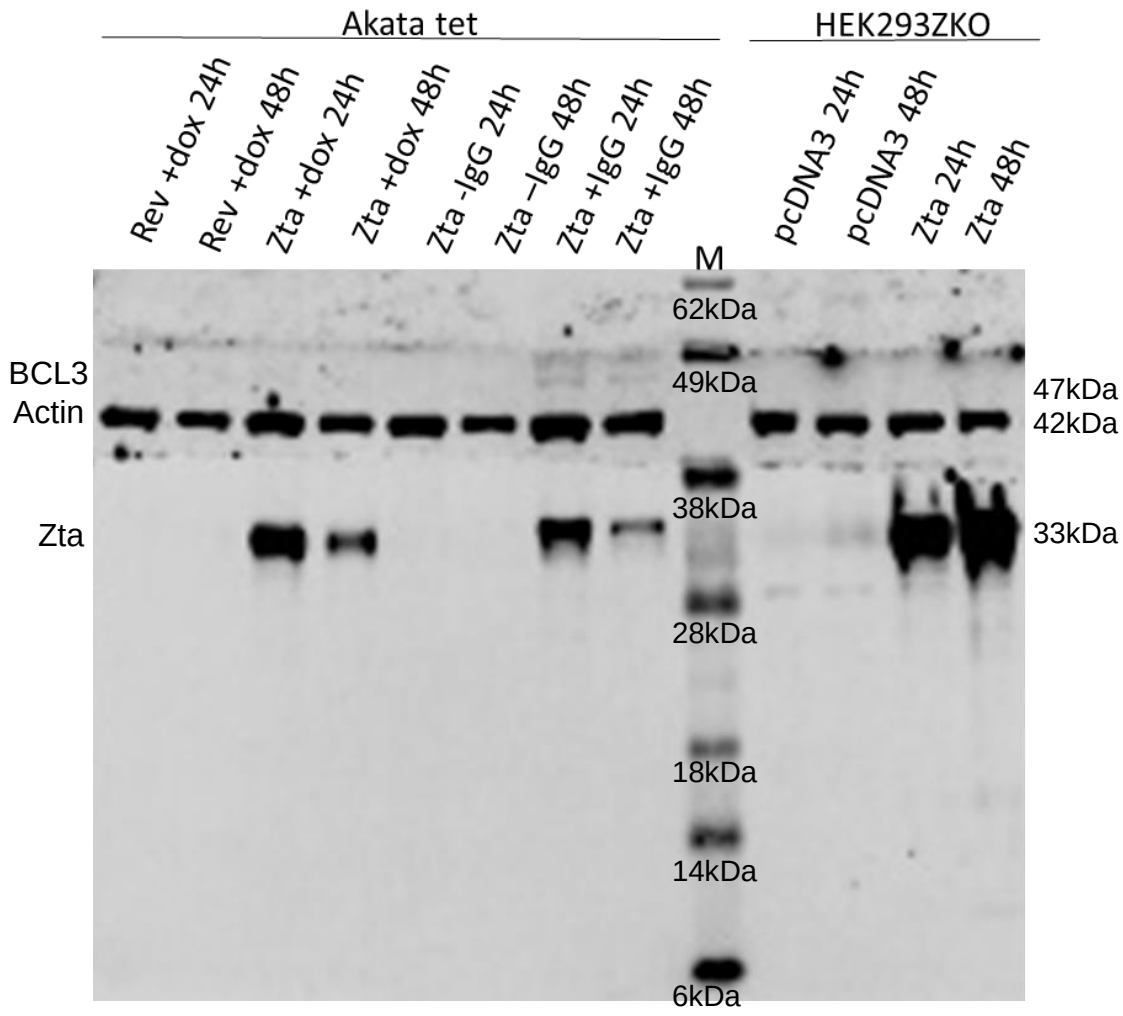


Figure 53: Western Blot showing the expression of Zta, host cell actin and the host transcription co-factor BCL3 in Akata or HEK 293ZKO protein extracts. Akata cells were un-induced or induced into lytic cycle with 0.125% IgG or 500 ng/ml doxycycline for the indicated time. HEK293 ZKO cell were induced into lytic cycle by transfection with Zta or not induced by transfection with a pcDNA3 vector. Blot was performed using the Santa Cruz BCL3 (c-14) antibody.

HEK293 ZKO cells was induced through transfection of a Zta expression plasmid. As in HONE1-EBV, the BCL3 protein levels appears to increase when EBV lytic cycle is induced with anti-IgG in Akata cells. The observed band is not based on

the IgG antibody used in Akata induction as the molecular weight of the rabbit F(ab')2 fragment antibody is approximately 110kDa (supplier information) and the BCL3 band is present just below the 49kDa marker. This is consistent with the BCL3 band seen in cells not induced with anti-IgG (e.g. HONE1-EBV).

The increase of BCL3 in Akata cells containing EBV induced into lytic cycle with anti-IgG in comparison to untreated cells was confirmed. The band intensity and corresponding amounts of Zta and BCL3 normalised to actin levels were quantified (Figure 54).

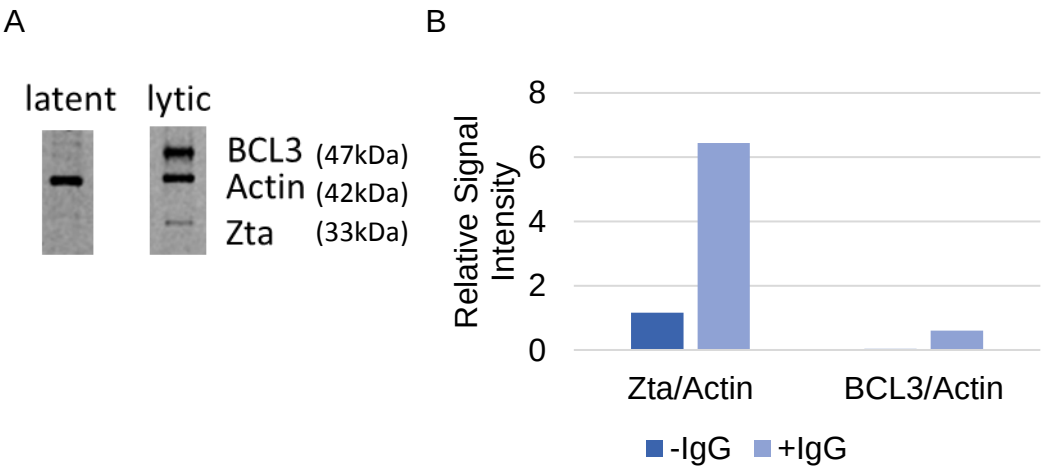


Figure 54: Western Blot (A) and band quantification (B) showing the expression of Zta, host cell actin and the host transcription factor BCL3 in Akata cells. Akata cells were induced into lytic cycle with 0.125% IgG for 24 hours, or not induced. Blot was performed using the Santa Cruz (c-14) x antibody.

As in HONE1-EBV cells, the amount of BCL3, p50 and p52 RNA, corresponding to the increase in BCL3 protein level, was investigated in Akata cells. This confirmed the absence on an increase in BCL3 transcription (Figure 55). However, as opposed to HONE1-EBV, the level of p50 RNA does not increase. This may suggest an epithelial-specific response of p50 to EBV lytic cycle.

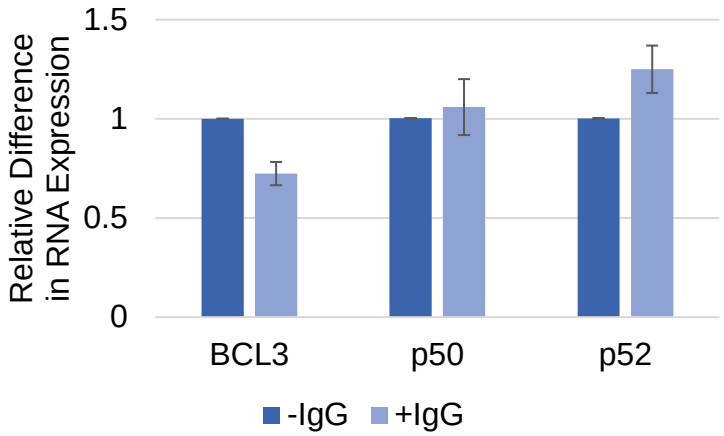


Figure 55: qPCR result showing the amounts of BCL3, p50 and p52 RNA in Akata cells. The lytic cycle was induced with 0.125% IgG for 24 hours, or unindicted. Values are represented as mean ±95% Confidence Interval.

Previous results, including MACS peak intersection of BCL3 and Zta peaks, have largely focussed on the human genome. Arvey *et al*. have mapped the binding of BCL3 in latently infected B95-8 LCLs to the EBV genome [25]. This shows a clear binding peak for BCL3 at the BHLF1 promoter (OriLyt). The western blot in Figure 56A shows the expression of BCL3 in reverse cross-linked chromatin of Akata cells containing lytic EBV. The results of a corresponding ChIP qPCR experiment for Zta, BCL3 and a negative control IgG antibody are shown in Figure 56B. The

enrichment for OriLyt is significantly higher when immunoprecipitation was performed with an antibody specific for Zta or BCL3 in comparison to a non-specific IgG control. This suggests that BCL3 binds to OriLyt in Akata cells and that overlapping binding sites for BCL3 and Zta can be extended to the viral genome.

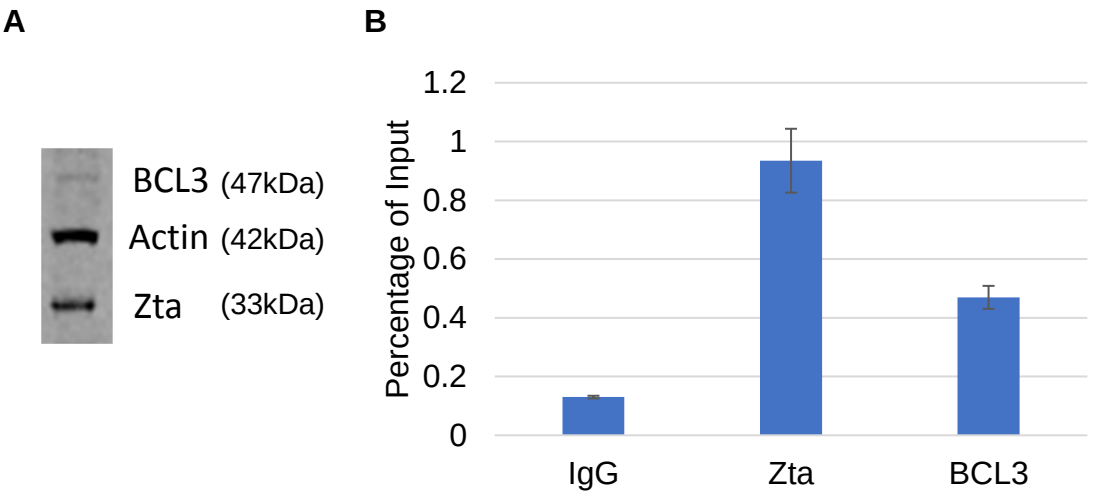


Figure 56: BCL3 association with OriLyt in Akata cells. Western Blot (A) showing the expression of Zta, host cell actin and the host transcription factor BCL3 in Akata cells. Blot shows reverse cross-linked chromatin derived from Akata cells induced into lytic cycle with 0.125% IgG for 24 hours. Blot was stained using the Santa Cruz BCL3 (c-14) x antibody. The corresponding ChIP qPCR analysis (B) shows the amount of OriLyt bound by Zta, BCL3 or a negative control IgG antibody. The binding of Zta and BCL3 to OriLyt is significantly higher than that of IgG ($p<0.05$). Values are represented as mean ± SD.

5.3 Binding of BCL3 to the Human Genome

The previous section demonstrated that BCL3 is expressed in the same cell types in which the Zta ChIP sequencing experiments were performed when EBV is induced into lytic cycle. Based on these results and BCL3 ChIP sequencing results in GM12878 cells obtained from the ENCODE database, it was decided to investigate whether some of the BCL3 binding peaks correspond to BCL3 binding in the EBV immortalized B-lymphoblastoid cell line 3 (LCL3). These cells were derived from a person infected with the B95-8 strain of EBV [24]. LCL3 cells do not require induction of EBV into lytic cycle. In this section, results are presented which show binding of both, Zta and BCL3 to regions of binding peak overlaps on the human genome. The possibility of BCL3 and Zta binding to the same region of the human genome at the same time was further investigated using ChIP reChIP qPCR. This involved a first round of ChIP with a Zta-specific antibody followed by a second round of ChIP of the same sample with a BCL3-specific antibody. As is shown in this section, this approach yielded no evidence of co-binding.

An example of MACS peak overlaps for Zta in epithelial cells and Akata B cells, as well as BCL3 in GM12878 B cells at the promoter regions of THRA, SCIMP and BCL2A1 can be seen in Figure 57.

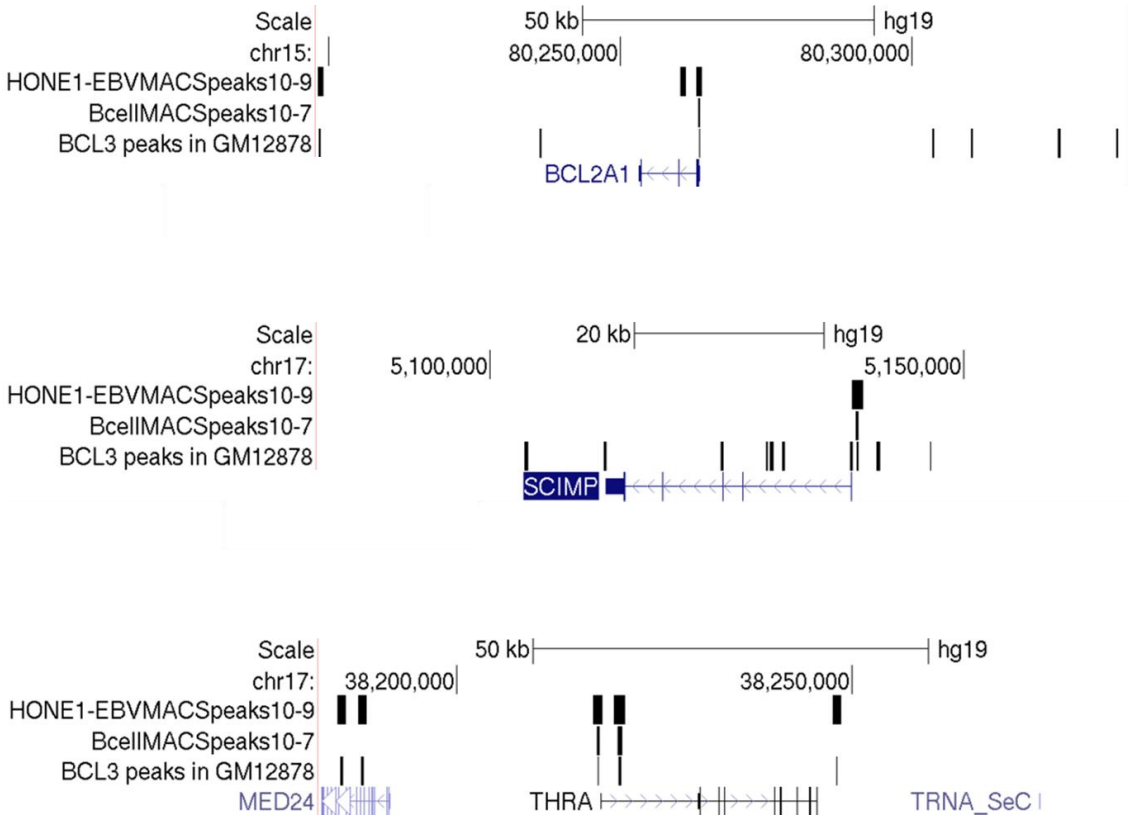


Figure 57: Example of ChIP sequencing MACS peaks around the BCL2A1, SCIMP and THRA gene. Peaks were visualised with the help of the UCSC genome browser.

The presence of BCL3, Zta and Actin in the sonicated LCL3 chromatin is presented in Figure 58.

Figure 59 shows that the percentage of the LCL3 cell population which are expessing Zta is relatively low, at ~2%. As stated in chapter 3, any effects of Zta, during experiments based on the average of the whole population, actually only represent a fraction of Zta positive cells (~2%). These experiments include the EBV copy number per cell, RNA levels, and western blots. ChIP qPCR experiments select BCL3 or Zta positive cells with the relevant antibody.

However, it does point to the necessity to increase the number of cells used for ChIP experimentsm, which are shown in the following section, in comparison to other cells lines which can be induced more efficiently.

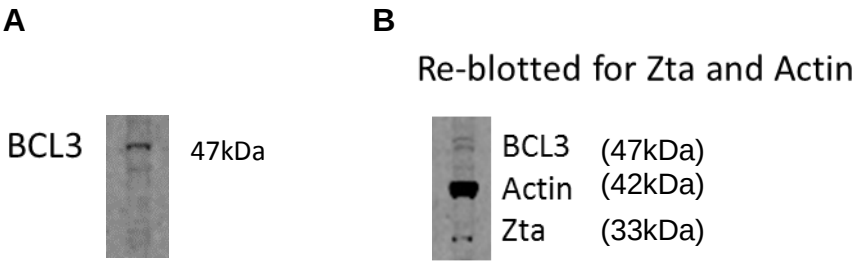


Figure 58: Western Blot showing the expression of BCL3 (A) in reverse cross-linked LCL3 chromatin together with a re-blot for Zta and host cell actin. The blot (A) was re-stained for the EBV early lytic cycle protein Zta and host cell actin (B). Blot was stained using the Santa Cruz BCL3 (c-14) antibody.

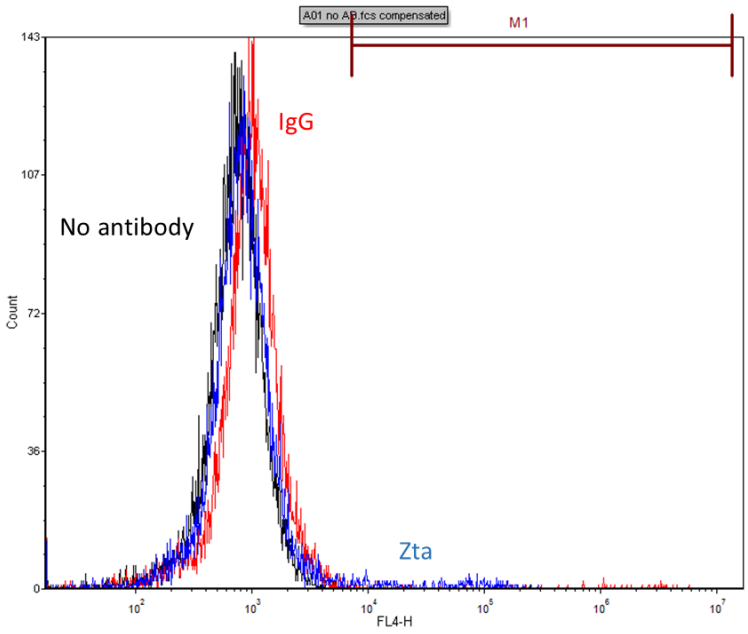


Figure 59: Overlay Histogram showing LCL3 cells containing Zta. Cells were not stained (black), stained with a negative control IgG antibody (red), showing background fluorescence, or Zta (blue).

Based on this overlap, a ChIP qPCR experiment was performed in LCL3 with primers specific for the peak overlap region or a flanking negative control region (Figure 60). No binding could be confirmed to the THRA peak region. This may be due to primer design. SCIMP and BCL2A1 are knows Zta binding sites and, as expected, a significantly higher amount of binding is observed for the peak region. The amount of binding to the BCL2A1 peak in comparison to its flanking region is also increased for BCL3.

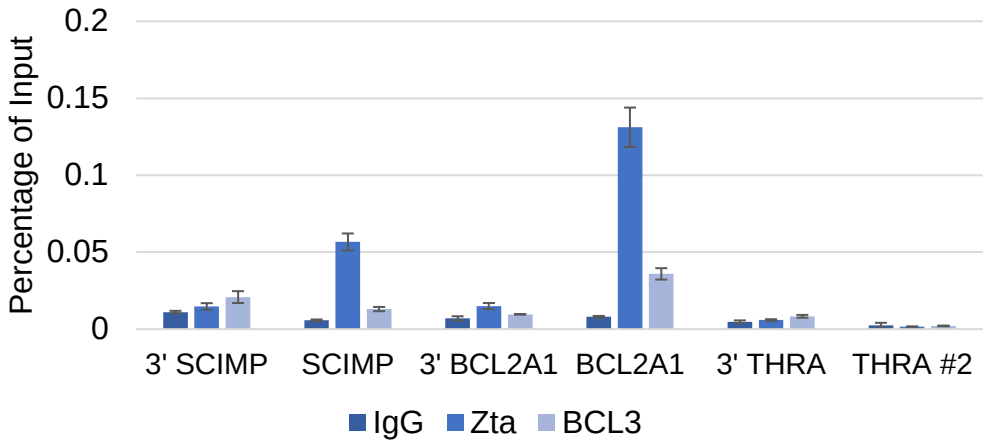


Figure 60: ChIP qPCR analysis showing the relative amount of LCL3 DNA bound by Zta, BCL3 or a negative control IgG antibody. Zta is known to bind to the promoter region of SCIMP and BCL2A1. These sites have previously been shown to be enriched during Zta ChIP experiments in comparison to 3' flanking regions. ChIP sequencing data analysis has revealed overlapping BCL3 MACS peaks in these regions of the human genome. The amount of ChIP DNA in the promoter region of BCL2A1, SCIMP and THRA peak two (THRA #2), and in corresponding flanking negative control regions, were quantified. The qPCR result shows significantly enhanced binding of Zta and BCL3 to the BCL2A1 peak in comparison to the flanking region ($p<0.05$). The increased binding of Zta, but not BCL3, to SCIMP is also significant ($p<0.05$). Values are represented as mean ± SD.

An intragenic binging peak for Zta and BCL3 in the EGFR gene has been visualised in Figure 61 below. In addition to probing for binding at gene promoters, this region was also examined in LCL3 (Figure 62). Both Zta and BCL3, but not negative control IgG, are shown to be enriched at this EGFR binding peak in comparison to the flanking region.

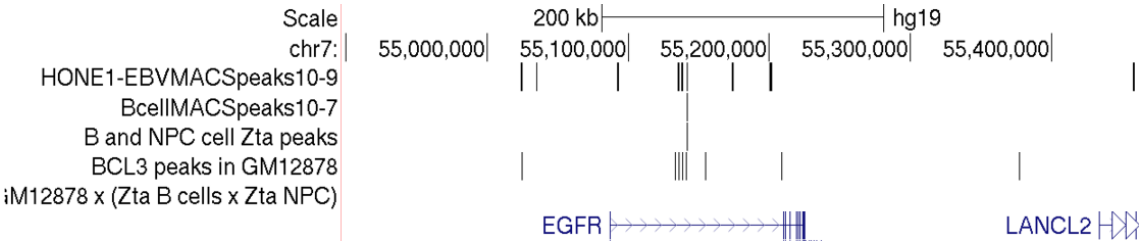


Figure 61: Example of ChIP sequencing reads around the EGFR gene. Zta and BCL3 binding site visualised, in the form MACS peaks, in the UCSC genome browser.

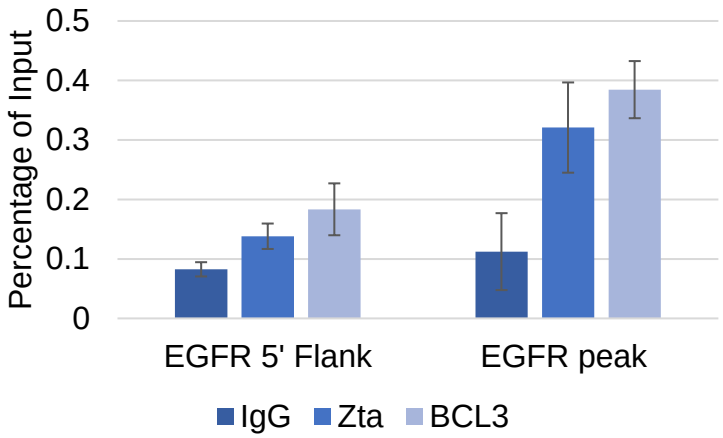


Figure 62: ChIP qPCR analysis showing the relative amount of LCL3 DNA bound by Zta, BCL3 or a negative control IgG antibody. EGFR is a new Zta binding region. This sites has not previously been shown to be enriched during Zta ChIP experiments in comparison to a 5' flanking region. ChIP sequencing data analysis has revealed overlapping BCL3 MACS peaks in this region of the human genome. The binding of Zta and BCL3 to EGFR peak is significantly higher than binding to the 5' Flanking region ($p<0.05$). Values are represented as mean ± SD.

To investigate if BCL3 and Zta are present at their common binding sites at the same time, a ChIP re-ChIP experiment was performed which involved a first round of ChIP with Zta or negative control IgG followed by a second round of ChIP for BCL3 and Zta (Figure 63). A very similar enrichment for binding to the BCL2A1 peak in comparison to the flank can be seen for samples immunoprecipitated with Zta followed by Zta or BCL3. This result suggests that Zta and BCL3 may both be present at the same binding site at the same time. Alternatively it is possible that Zta antibody was carried over after the first round of ChIP.

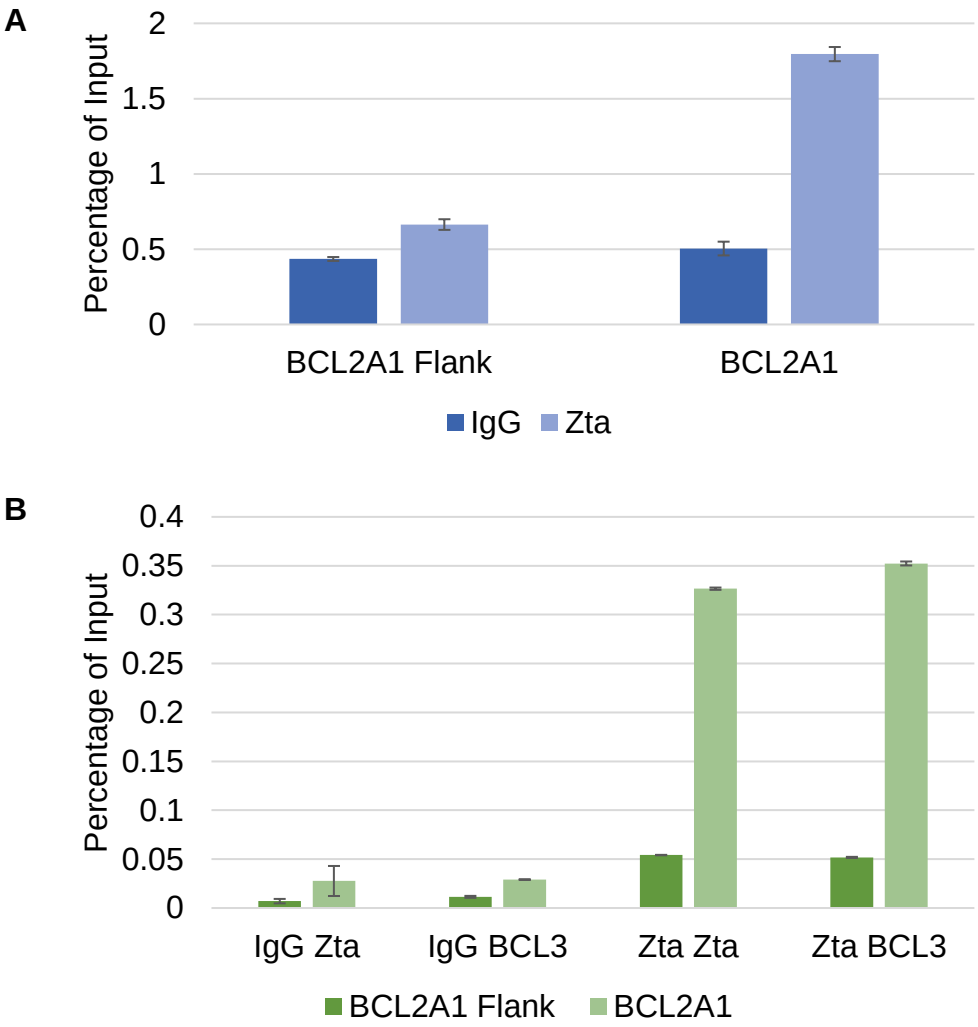


Figure 63: ChIP reChIP qPCR showing the relative amount of LCL3 DNA bound by Zta and a negative control IgG antibody (A) including a second round of ChIP for BCL3 and Zta (B). Zta is known to bind to the promoter region of BCL2A1. This site has previously been shown to be enriched during Zta ChIP experiments in comparison to 3' flanking regions in this cell line. ChIP sequencing data analysis has revealed overlapping BCL3 MACS peaks in this regions of the human genome. The amount of ChIP DNA in the promoter region of BCL2A1 and in its corresponding flanking negative control region, were quantified. The binding of Zta to the BCL2A1 peak is significantly higher than to the flanking region ($p<0.05$). Values are represented as mean ± SD.

It was decided to follow this result up further with the aim to perform a ChIP re-ChIP sequencing experiment for sites bound by Zta and BCL3. For this purpose, a large scale chromatin preparation was performed. The quality of this chromatin was checked by southern blot and qPCR. The resulting fragment size can be seen in Figure 64.

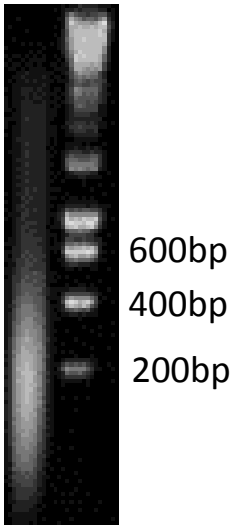


Figure 64: Agarose Gel showing the size, in base pairs (bp), of sonicated LCL3 chromatin fragments. The ideal fragment size is 100-300bp

In addition to the control first round IgG ChIP, a first round Zta, followed by a second round IgG ChIP was performed as a control to confirm whether or not BCL3 bind together with Zta. Figure 65 shows the qPCR results for the first round of ChIP with antibodies specific for BCL3, Zta or a negative control IgG with enhanced binding to the BCL2A1 peak in comparison to the flank. Following this first round of Zta ChIP, a second round of ChIP was performed using Zta as a positive control, IgG as a negative control, and BCL3 (Figure 66).

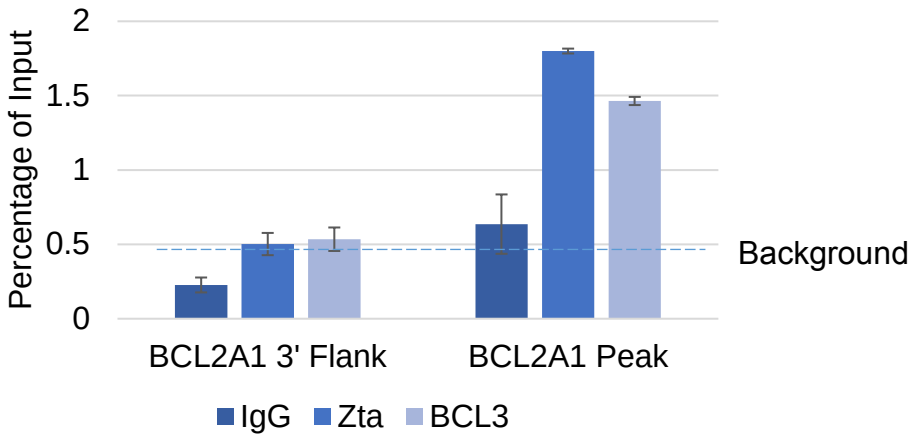


Figure 65: qPCR results of a ChIP experiment showing IgG, Zta and BCL3 binding to BCL2A1 in comparison to a 3' flanking region. This experiment was performed using the same antibodies on the same chromatin, at the same time in the same qPCR run as the ChIP reChIP experiment. Data represented as percentage of Input. The enhanced binding of Zta and BCL3, but not IgG, to the BCL2A1 peak in comparison to the flanking region is significant ($p<0.05$). Values are represented as mean ± SD.

The negative control re-ChIP with IgG appears to show the same enrichment as the re-ChIP with BCL3 (Figure 66). As IgG has been shown to be a suitable negative control in previous ChIP qPCR experiments, this result suggests that BCL3 and Zta may not bind to the promoter region of BCL2A1 together, but rather that the enrichment of this region is caused by carry-over of the Zta antibody from the first-round of ChIP. To investigate the possibility that the enrichment of binding to the peak in comparison to the flank after the first round of Zta ChIP is based on inefficient removal of the first round antibody, a ChIP was performed using Zta antibody denatured in the same way as after the first round of ChIP.

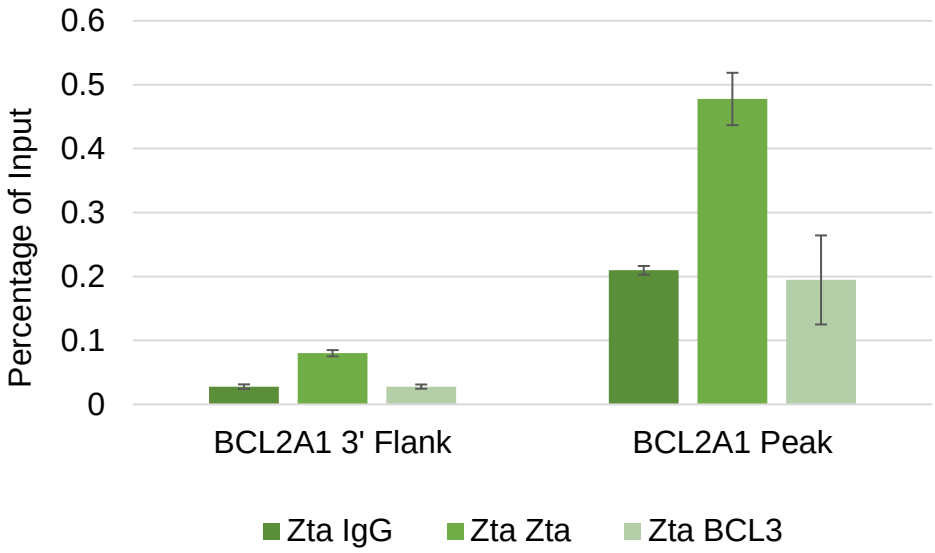


Figure 66: ChIP reChIP qPCR analysis showing the relative amount of LCL3 Zta ChIP DNA also bound by IgG, BCL3 or Zta. This site has previously been shown to be enriched during Zta and BCL3 ChIP experiments in comparison to 3' flanking regions in this cell line. ChIP sequencing data analysis has revealed overlapping Zta and BCL3 MACS peaks in this regions of the human genome. Following a first round of Zta ChIP, the increase in DNA precipitated by Zta and IgG, but not BCL3, to the BCL2A1 peak in comparison to the flanking region is significant ($p<0.05$). Data represented as percentage of input. Values are represented as mean ± SD.

As can be seen in Figure 67, a reduced, but significant enrichment of the BCL2A1 promoter region is still achieved when performing ChIP using the denatured Zta antibody. The fold change of binding to the peak in comparison to the flanking region is very similar to the previously shown second round of ChIP with IgG or BCL3 (Figure 66). If BCL3 generally bound to the same regions as Zta, a higher

enrichment, similar to that seen for a second round Zta ChIP, would be expected. These results demonstrate that a Zta ChIP, BCL3 re-ChIP sequencing experiment is unlikely to be valuable.

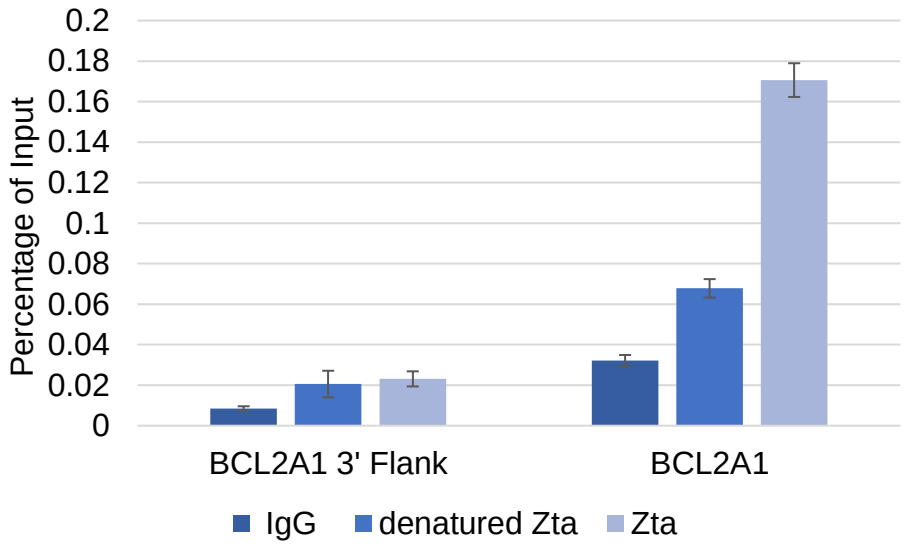


Figure 67: qPCR results of ChIP experiment showing the binding of a negative control IgG antibody, Zta and denatured Zta to BCL2A1 in comparison to a 3' flanking region. The enhanced binding to the peak in comparison to the flanking region of BCL2A1 is significant in all cases ($p<0.05$). Data represented as percentage of Input.Values are represented as mean ± SD.

5.4 BCL3 Protein Levels

To investigate the upregulation of BCL3 as a result of lytic cycle induction more closely, FACS was performed to determine the percentage of BCL3 positive cells. Figure 68 shows the amount of Akata cells positive for BCL3 or a negative control antibody. The BCL3 antibody seems to have a higher amount of background staining. The amounts of both control and BCL3 positive cells appear to increase when lytic cycle is induced. The induction process requires a rabbit IgG antibody

and, both, control and BCL3, antibodies used for FACS are also of rabbit origin. To detect them, a rabbit specific, fluorescent antibody is used which leads to false positive results in induced cells.

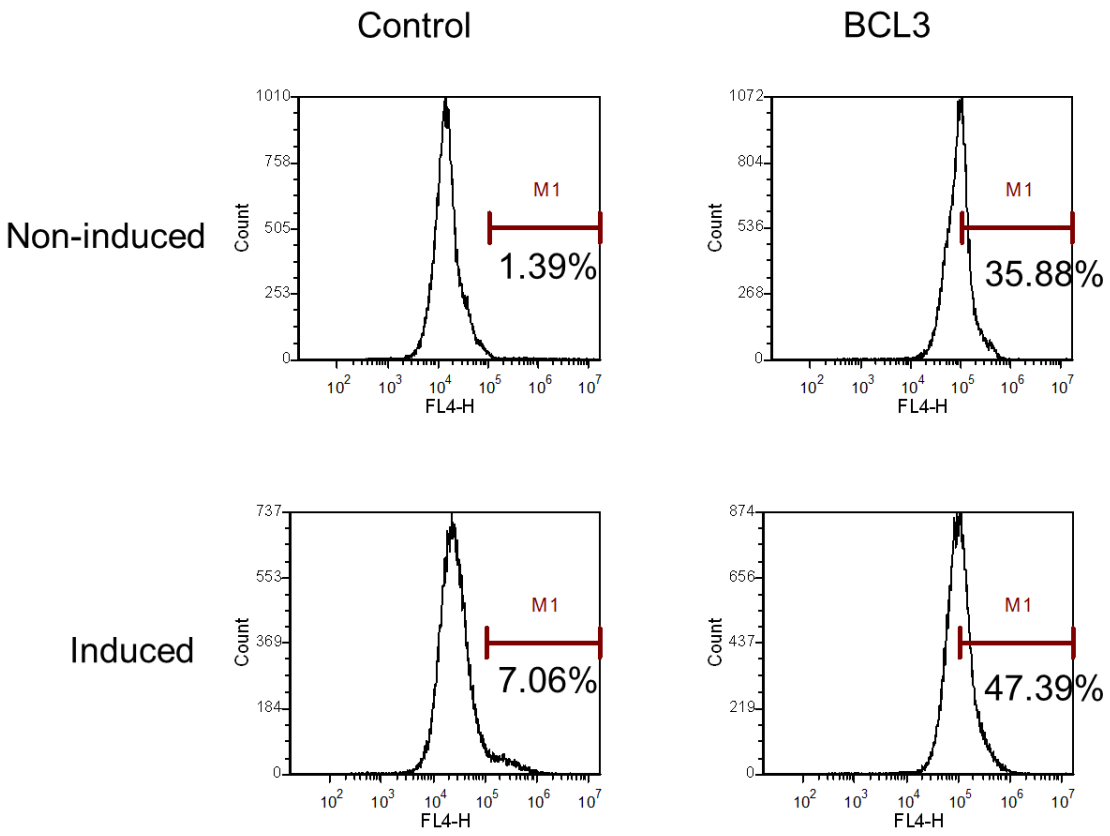


Figure 68: Akata FACS showing percentage of BCL3 positive cells. Akata cells were not induced (top row) or induced with 0.125% rabbit IgG for 48 hours (bottom row). Cells were fixed and stained with rabbit negative control (left column) or rabbit BCL3 antibodies (right column) followed by a rabbit specific fluorescent secondary antibody.

Overlay histograms were generated to directly compare the control to the BCL3 specific antibody (Figure 69). In conjunction with BCL3 staining, the amount of viral capsid antigen (VCA) positive cells was also investigated. This protein is

used as a marker of cells containing EBV in late lytic cycle. This reveals that about 5% of induced cells are VCA positive (Figure 70). For direct comparison of VCA levels in induced and non-induced cells, Figure 71 shows these results in an overlay histogram. As opposed to BCL3, the VCA antibody and control are of mouse origin, so the induction of lytic cycle with rabbit IgG should not directly affect the result.

As the percentage of VCA positive cells only represents the amount of cells containing EBV in late lytic cycle, FACS was repeated including a Zta (BZ1) specific antibody for cells containing EBV in early lytic cycle. In addition the lytic cycle was induced using a doxycycline inducible plasmid containing Zta, or a negative control reverse plasmid, instead of the rabbit IgG antibody, to prevent false positive staining. Induction was undertaken for both 24 and 48 hours. Table 11 summarises the percentage of cells positive for the indicated targets. This reveals a greater increase of BCL3 amount (~15%) in cells containing EBV induced for 48 hours and thus in more established lytic cycle in comparison to 24 hour induction. VCA staining did not increased with longer induction. This may be due to cell lysis. The amount of cells containing EBV in early lytic cycle (BZ1) is shown to have almost halved at the longer induction time.

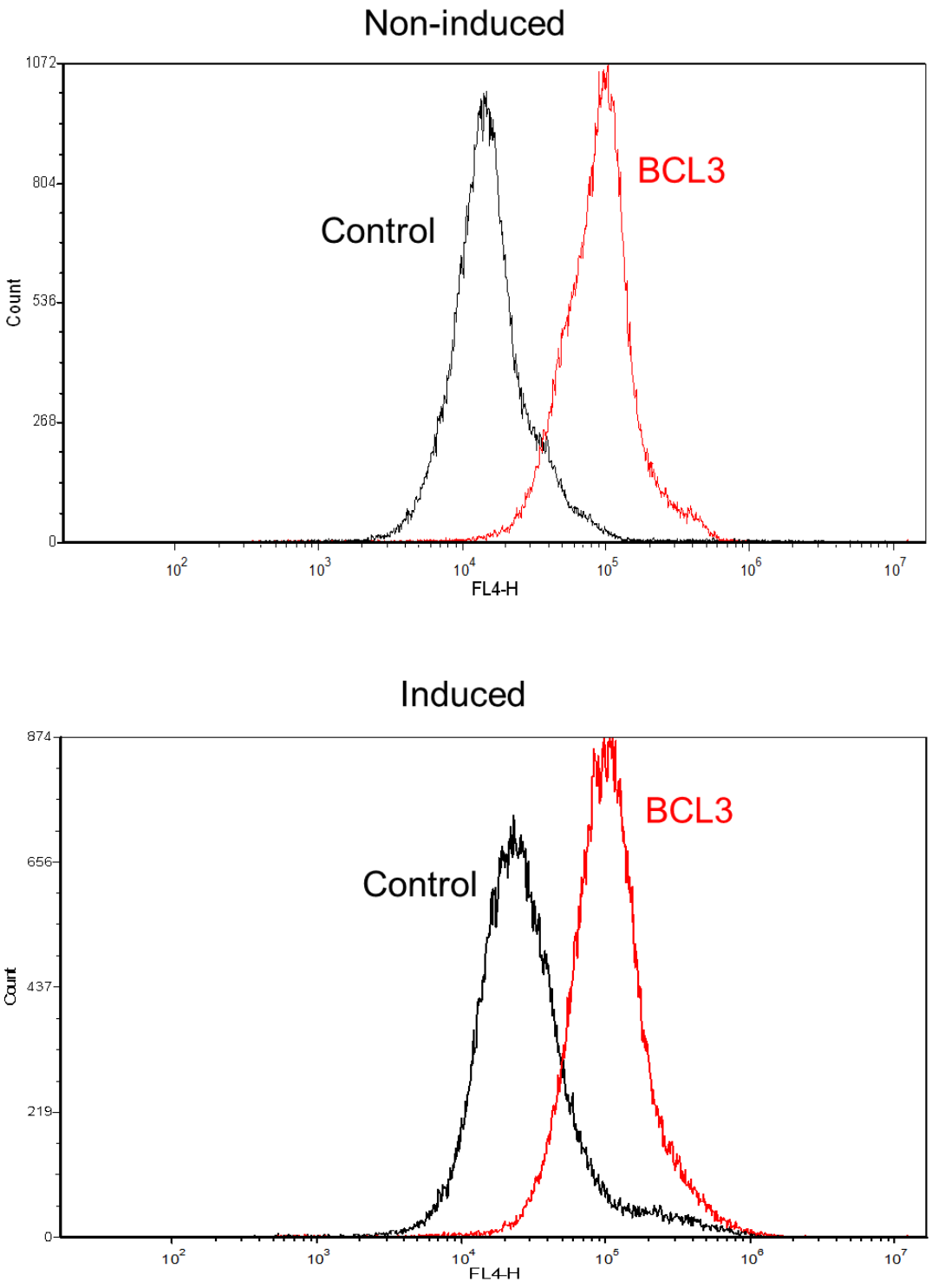


Figure 69: FACS overlay histograms showing the amount of BCL3 positive cells. Akata cells were not induced (top) or induced (bottom) with 0.125% rabbit IgG for 48hours. Cells were fixed and stained with a control (black) or BCL3 antibody (red) followed by a specific fluorescent secondary antibody.

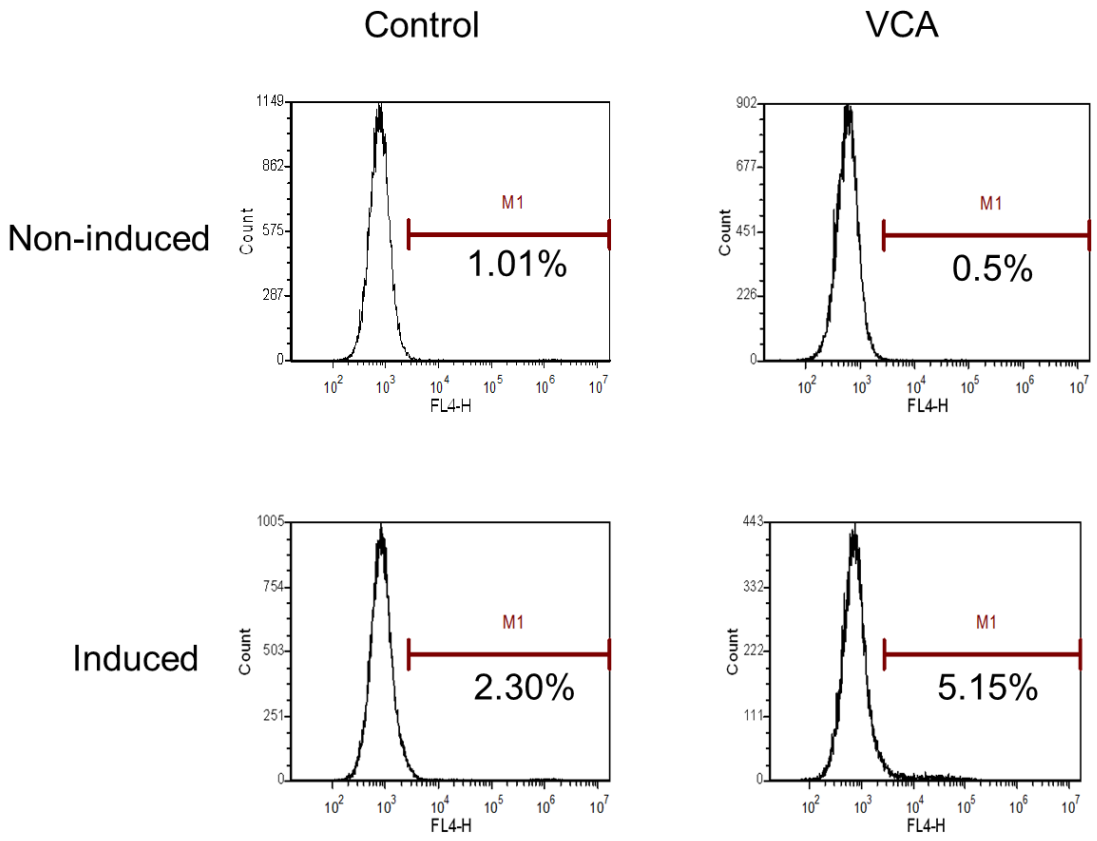


Figure 70: Akata FACS showing percentage of VCA positive cells. Akata cells were not induced (top row) or induced with 0.125% rabbit IgG for 48 hours (bottom row). Cells were fixed and stained with mouse negative control (left column) or mouse VCA antibodies (right column) followed by a mouse specific fluorescent secondary antibody.

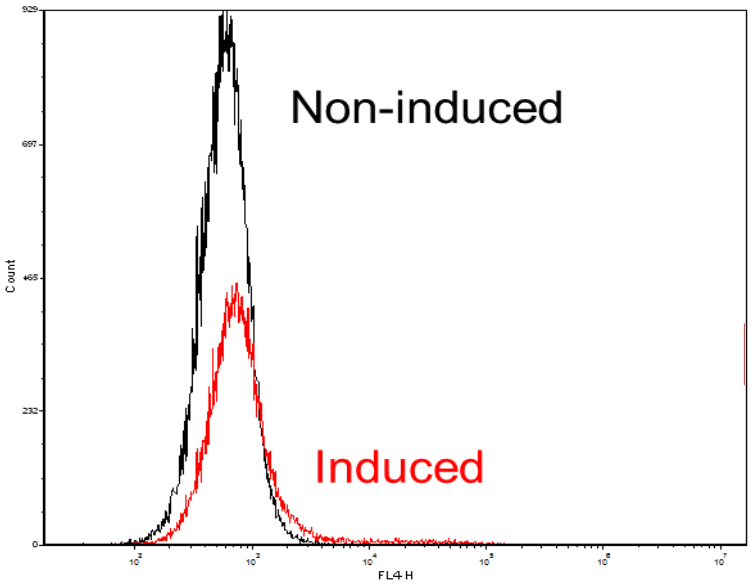


Figure 71: FACS overlay histogram showing the amount of VCA positive cells. Akata cells were not induced (black) or induced with 0.125% rabbit IgG for 48 hours (red). Cells were fixed and stained with a mouse VCA antibody followed by a mouse specific fluorescent secondary antibody.

Target	Non-induced		Induced	
	24hrs	48hrs	24hrs	48hrs
Mouse Control	1.06	1.08	0.80	1.32
Rabbit Control	1.10	1.04	2.33	1.93
BCL3 (Rabbit)	76.11	47.04	83.64	62.15
VCA (Mouse)	0.28	0.44	7.13	6.34
BZ1 (Mouse)	0.35	0.18	10.99	5.65

Table11: Table showing the percentage of Akata cells positive for a specific FACS antibody target. Akata cells containing either a Zta doxycycline inducible plasmid or a negative control reverse plasmid were treated with 500 ng/ml doxycycline for 24hrs or 48hrs. Cells were fixed and stained with a BCL3, VCA, BZ1 or control antibody followed by a specific fluorescent secondary antibody.

The previous experiment was repeated using Akata cells induced with rabbit IgG antibody instead of the Zta doxycycline inducible plasmid (Table 12). This shows that there is a substantial increase in the amount of cells stained with negative control, rabbit-specific secondary antibody when cells are induced with rabbit IgG, especially at 48 hours of induction (46.81%). The amount of BCL3 cannot be seen to increase during this experiment. As before, the amount of Zta (BZ1) is halved at 48 hours of induction in comparison to 24 hours. However, surprisingly staining for VCA appears higher at 24 hours. This might be because induction with IgG utilises a different mechanism to induction with doxycycline which leads to a faster progression to late lytic cycle.

Target	Non-induced		Induced	
	24hrs	48hrs	24hrs	48hrs
Mouse Control	1.39	1.02	0.40	0.93
Rabbit Control	0.92	3.04	11.51	46.81
BCL3 (Rabbit)	46.58	88.22	45.03	82.07
VCA (Mouse)	0.36	0.38	8.23	0.57
BZ1 (Mouse)	0.37	0.56	4.75	2.45

Table 12: Table showing the percentage of Akata cells positive for a specific FACS antibody target. Akata tet cells were not induced or induced into lytic cycle with 0.125% rabbit IgG for 24hrs or 48hrs. Cells were fixed and stained with a BCL3, VCA, BZ1 or control antibody followed by a specific fluorescent secondary antibody.

To overcome the problem of false positive staining of the rabbit specific secondary antibody in Akata cells induced with rabbit IgG, a new fluorescent hamster BCL3 antibody with a hamster control florescent IgG antibody were used for FACS. As can be seen in Figure 72, this combination does not show the presence of BCL3 in Akata cells containing induced or non-induced EBV above the level of negative control staining. The inability to show an increase in BCL3 levels as a result of lytic cycle induction using this method can be clearly seen in the overlay histogram in Figure 73. The corresponding amount of VCA positive cells is shown in Table 13, demonstrating that induction was successful. The Table also shows that it was not possible to detect any cells positive for both, BCL3 and VCA.

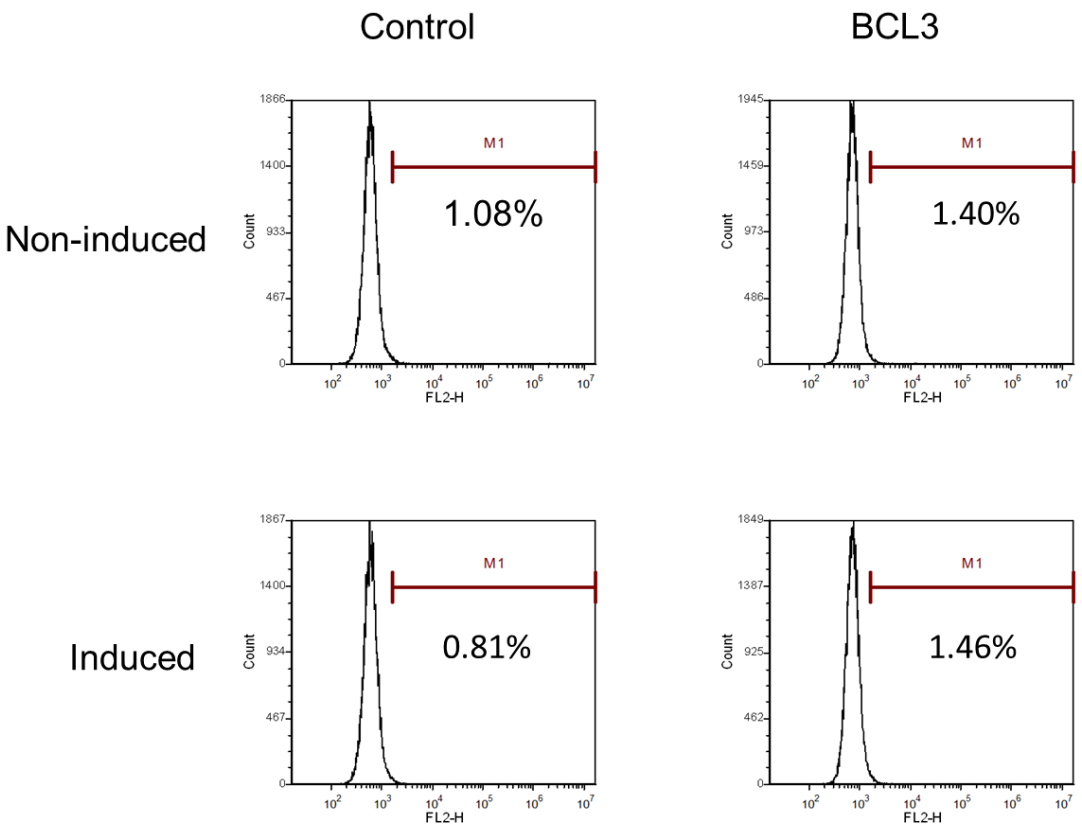


Figure 72: Histograms showing the percentage of Akata tet cells positive or negative for BCL3 expression. Cells were not induced (top row) or induced (bottom row) into lytic cycle with 0.125% rabbit IgG for 24hrs, fixed and stained with a fluorescent control (left column) or fluorescent hamster BCL3 antibody (right column).

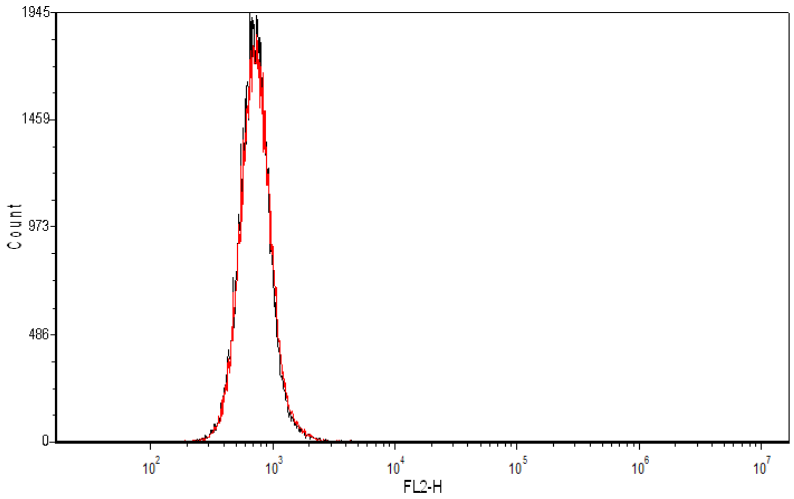


Figure 73: Overlay-histogram showing the expression of BCL3 in Akata tet cells non-induced (black) or induced (red) into lytic cycle. Induction was performed with 0.125% rabbit IgG for 24hrs. Cells were fixed and stained with a fluorescent hamster BCL3 antibody. Figure shows that there is no evidence by FACS of BCL3 upregulation in Akata cells.

	Non-induced	**Induced**
Mouse Control	1.08	0.89
VCA (Mouse)	0.12	11.79
VCA (Mouse) and BCL3 (Hamster)	0.09	0.16

Table 13: Table showing the percentage of VCA and "BCL3 and VCA" positive Akata tet cells corresponding to the above BCL3 amounts. Cells were induced into lytic cycle with 0.125% rabbit IgG for 24hrs, fixed and double stained with a fluorescent rabbit BCL3 and a mouse VCA antibody followed by a mouse fluorescent secondary antibodies.

5.5 BCL3 Knock-Down

To investigate the relevance of BCL3 for EBV lytic cycle reactivation, it was decided to use BCL3 specific siRNA to knock down its expression. To control for the transfection of siRNA, a Cy3 stained control siRNA was included in the transfection at a ratio of 3:1. Initial trials showed that, 72 hours after transfection with 750nM siRNA, 32.59% of cells are positive for the Cy3 marker (Figure 74). To optimise the transfection protocol, different concentrations of siRNA were used and the subsequent transfection efficiency was investigated with the help of the Cy3 marker siRNA.

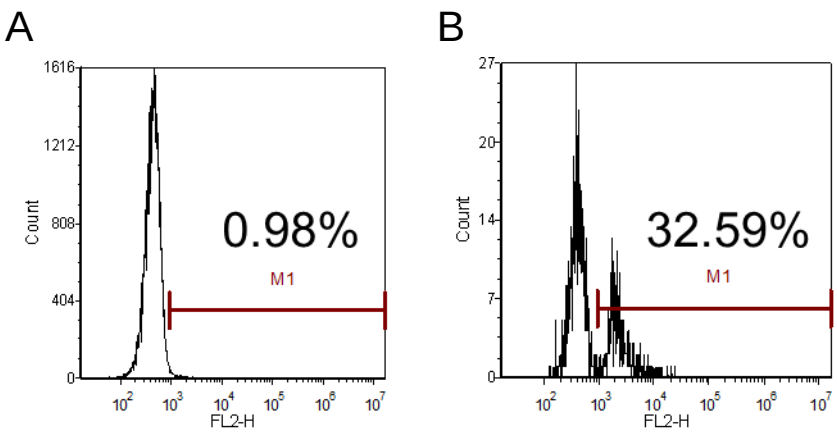


Figure 74: FACS histograms showing the percentage of siRNA transfected cells. Using the NEON, Akata cells were not transfected (A) or transfected (B) with 750nM siRNA at a 3:1 ratio of control siRNA to Cy3 stained siRNA. 72hrs after transfection cells were fixed and FACS was performed to show the amount of Cy3 positive cells. Cells that were not transfected show a background fluorescence of 0.98%. Transfected cells are 32.59% Cy3 positive, as detected in the FL2 channel.

As can be seen in Figure 75, 24 hours after transfection, of the three tested concentrations, 200nM siRNA resulted in the highest amount of siRNA positive cells at 87.76%. Transfection with 750nM resulted in a similar percentage of Cy3 positive cells as shown in the previous Figure (35.23%).The lowest amount of siRNA positive cells was achieved with the highest concentration of siRNA at 5µM.

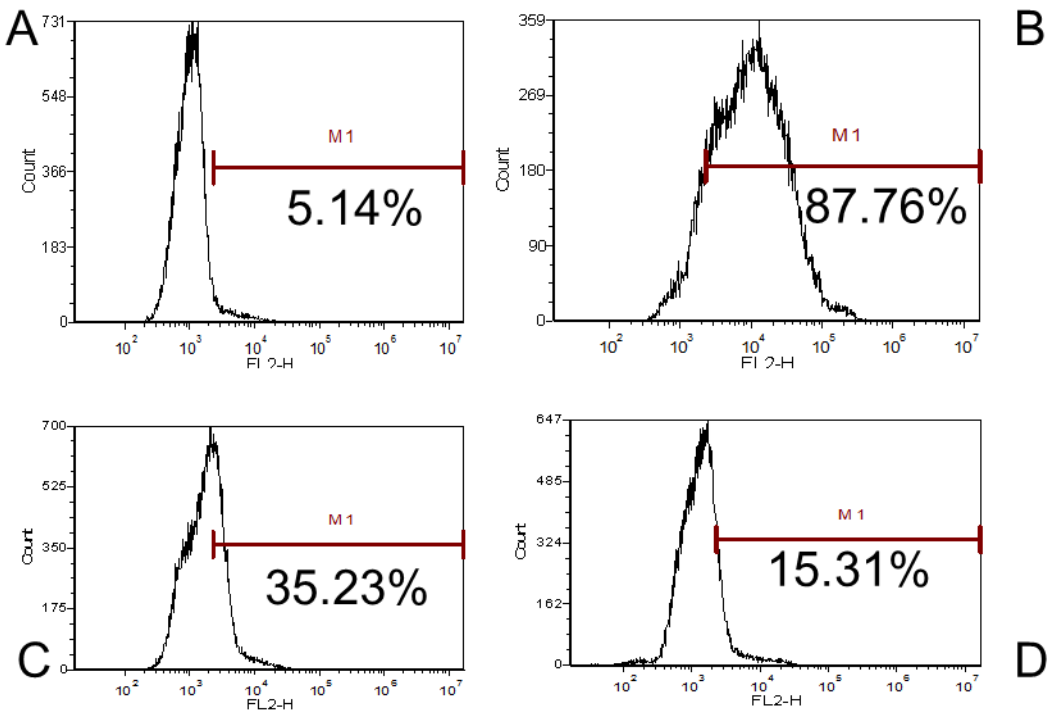


Figure 75: FACS histograms showing the percentage of siRNA transfected cells. Akata cells were not transfected (A) or transfected with 200nM (B), 750nM (C) or 5µM (D) siRNA at a 3:1 ratio of control siRNA to Cy3 stained siRNA. 24hrs after transfection cells were fixed and FACS was performed to show the amount of Cy3 positive cells and demonstrate transfection efficiency at different concentrations.

In addition to testing for cells transfected with siRNA 24 hours after transfection, the amount of cells positive for the Cy3 siRNA marker was also quantified 72 hours after transfection. The results of this experiment are summaries in Figure 76 below, showing that, by this point, no substantial amount of siRNA can be detected in the Akata cells.

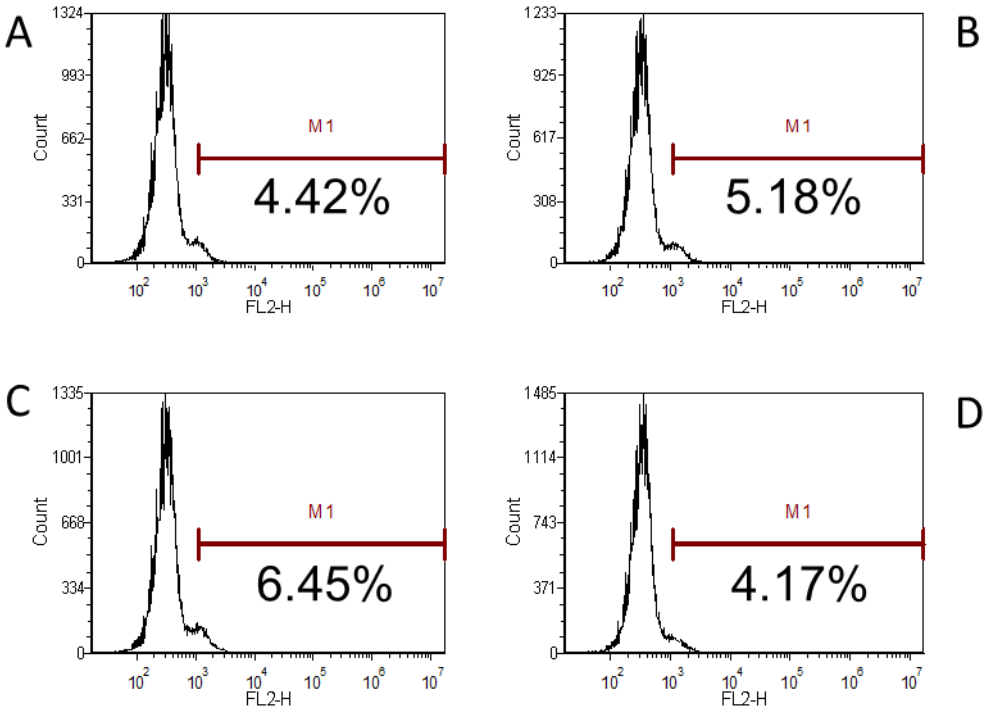


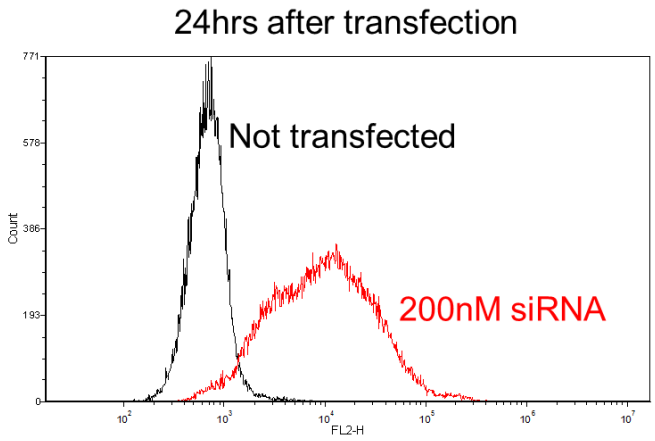
Figure 76: FACS histograms showing the percentage of siRNA transfected cells. Akata cells were not transfected (A) or transfected with 200nM (B), 750nM (C) or 5μM (D) siRNA at a 3:1 ratio of control siRNA to Cy3 stained siRNA. 72hrs after transfection cells were fixed and FACS was performed to show the amount of Cy3 positive cells and demonstrate transfection efficiency at different concentrations.

The background fluorescence in the FL2 channel is at 4.42% (Figure 76 A). The highest signal in transfected cells is 2.03% above the background for cells transfected with 750nM siRNA. In contrast, the highest amount of positive cells transfected with 200nM siRNA by 24 hours was 82.62%.

Figure 77 shows a comparison of the amount of cells positive for the Cy3 siRNA marker 24 or 72 hours after transfection in two overlay histograms. This clearly shows that the majority of cells have taken up the siRNA 24 hours after transfection but loose it over time.

Unfortunately it was not possible to repeat this result. This may be due to a problem with the transfection, the Cy3 siRNA itself or the FACS procedure. To investigate if the problem could be resolved by not fixing the cells during FACS, Akata cells were transfected with 0, 200nM or 5µM siRNA and either fixed or not fixed before FACS was performed. Results of this experiment are shown in Figure 78. No clear difference can be observed for transfection with 200nM and only a slight difference can be seen when transfection was performed with 5µM. The amount of Cy3 siRNA is 25 times higher during the transfection with 5µM in comparison to 200nM. As transfection was previously shown to be more efficient with 200nM siRNA, this result could point to a problem with the fluorescent siRNA marker.

A



B

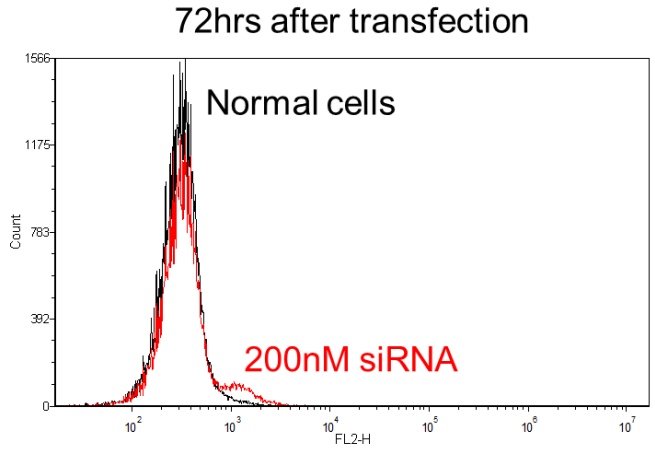


Figure 77: Overlay histograms showing the amount of cells positive for Cy3 siRNA marker in Akata cells. Cells were not transfected (black) or transfected with 200nM (red) siRNA at a 3:1 ratio of control siRNA to Cy3 stained siRNA. Cells were fixed and FACS was performed to show the amount of Cy3 positive cells and demonstrate transfection efficiency 24 hour (A) or 72 hours (B) after transfection.

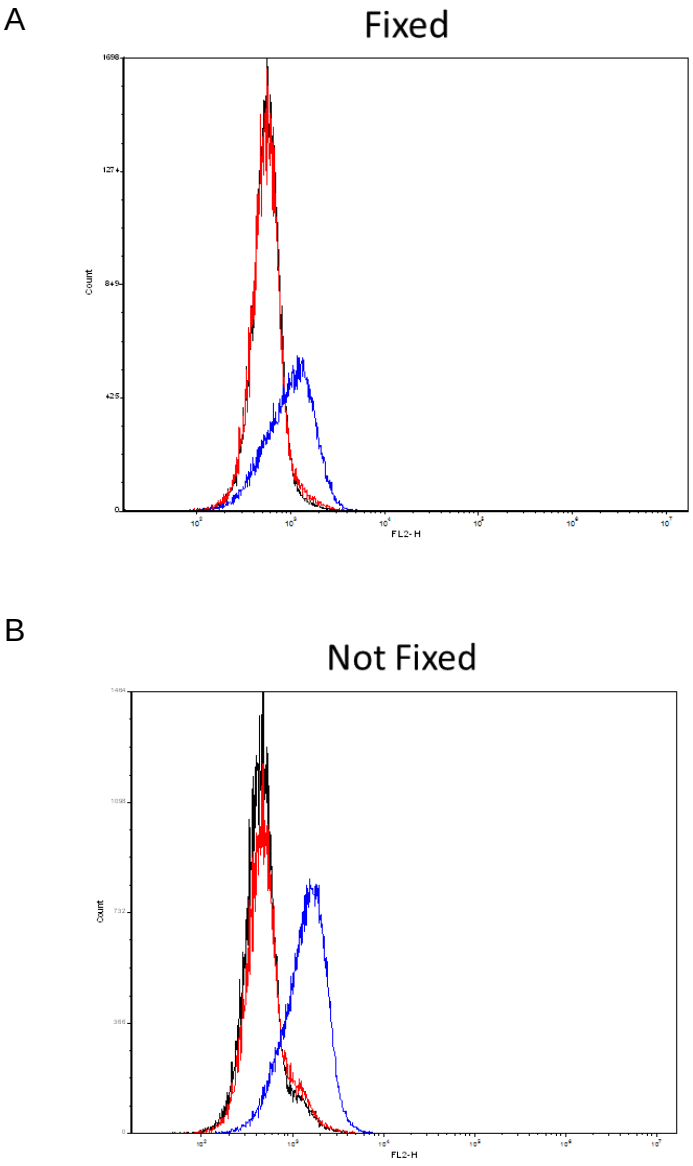


Figure 78: Overlay-histogram showing the amount of the siRNA marker Cy3. Akata tet cells were transfected with no siRNA (black), 200nM (red) or 5µM (blue) of negative control siRNA at a 3:1 ratio of siRNA to Cy3 stained siRNA. Cells were fixed (A) or not fixed (B) before FACS was performed 24 hours after transfection. Lytic cycle was not induced.

To investigate whether the knock-down of BCL3 has an impact on the switch from EBV latency to lytic cycle, Akata cells were transfected with 200nM siRNA, followed by the induction of lytic cycle. Based on the results shows in the previous Figures, induction of lytic cycle was undertaken 24 hours after siRNA transfection, at which point the highest amount of siRNA was detected in the cells. Three different BCL3 siRNAs were used in the attempt to knock down BCL3. No significant difference can be seen between EBV loads in cells treated with control siRNA in comparison to siRNA1-3 (Figure 79). As cells were harvested 24 hours after induction of lytic cycle, it is likely that this result is, at least partly, due to the fact that lytic cycle has not progressed far enough at this time.

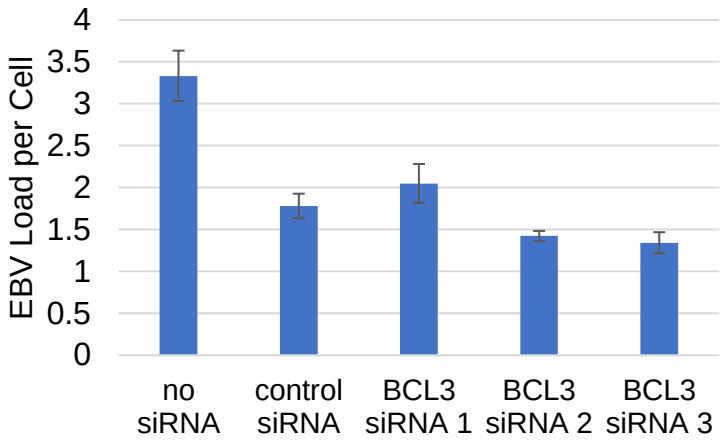


Figure 79: qPCR results showing the relative amount of EBV genome copies per human genome numbers. Akata tet cells were transfected with 200nM of the indicated siRNA. Lytic cycle was induced 24hrs after transfection with 0.125% rabbit IgG for 24hrs. There is no significant difference between EBV loads in cells treated with control siRNA in comparison to siRNA1-3. Averages are based on triplicate pipetting repeats. Values are represented as mean ± SD.

The amount of BCL3 RNA in Akata cells 24 hours after BCL3 siRNA knock down using 200nM siRNA were investigated with the help of a BCL3 revers transcribed cDNA qPCR (Figure 80). This revealed no significant decrease in BCL3 RNA levels of cells transfected with no, or non-specific, siRNA in comparison to BCL3 siRNA. This result point to a possible problem with the siRNA. For this reason, new siRNA was used to investigate the effect of BCL3 knock-down on EBV's reactivation from latency.

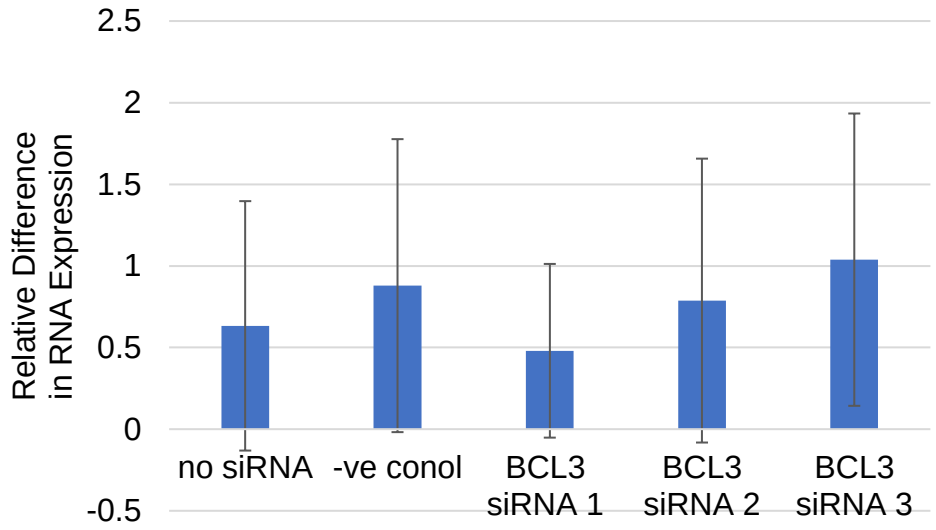


Figure 80: qPCR result showing the difference in BCL3 expression relative to BCL3 siRNA2. Akata cells were transfected, using the NEON, with 200nM of the indicated siRNAs for 24hrs. Values are represented as mean ± 95% Confidence Interval ($p>0.05$).

Figure 81 shows results obtained using the DharmaFECT transfection reagent and corresponding BCL3 siRNA pool to transfected HEK293 ZKO cells. The cells were also treated with ionomycin and phorbol 12-myristate 13-acetate (PMA) to

activate the B cell receptor signalling pathway. As shown in the previous section, the detection of BCL3 by western blot in HEK293 ZKO cells is poor. Nonetheless, when BCL3 is knocked-down in these cells, an increase of Zta protein abundance with a fold change of 1.4 can be observed 24 hours after hisZta expression vector transfection. It was shown in the previous section that BCL3 levels increase when Zta is expressed in lytic cycle. Based on these results, it is possible that Zta and BCL3 affect each other's protein levels.

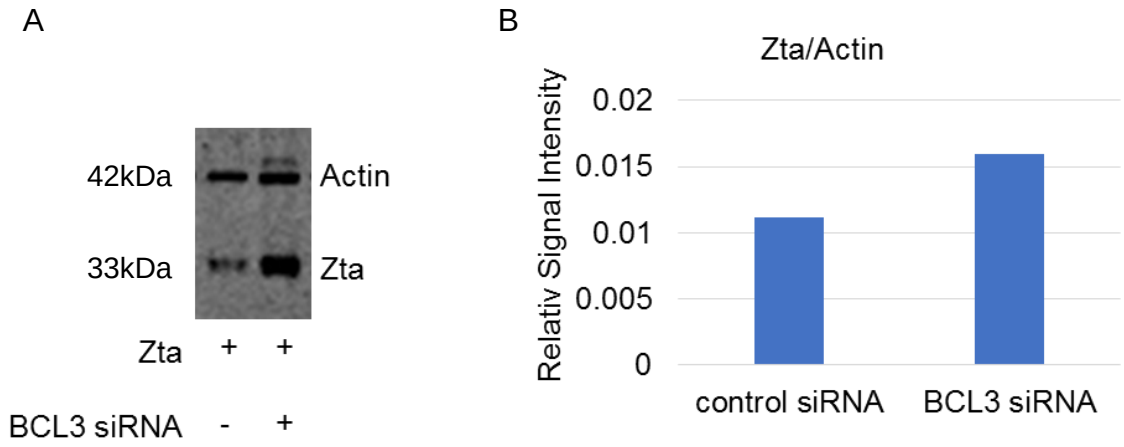


Figure 81: Western Blot (A) and corresponding band quantification (B) showing Zta and actin levels. HEK293 ZKO cells were transfected with a siRNA control or BCL3 specific siRNA followed by the hisZta expression vector and treatment with 1µM ionomycin and 20ng/ml PMA 24 hours later. Cells were harvested for western blotting a further 24 hours later. Blit was stained using the Santa Cruz BCL3 (c-14) antibody.

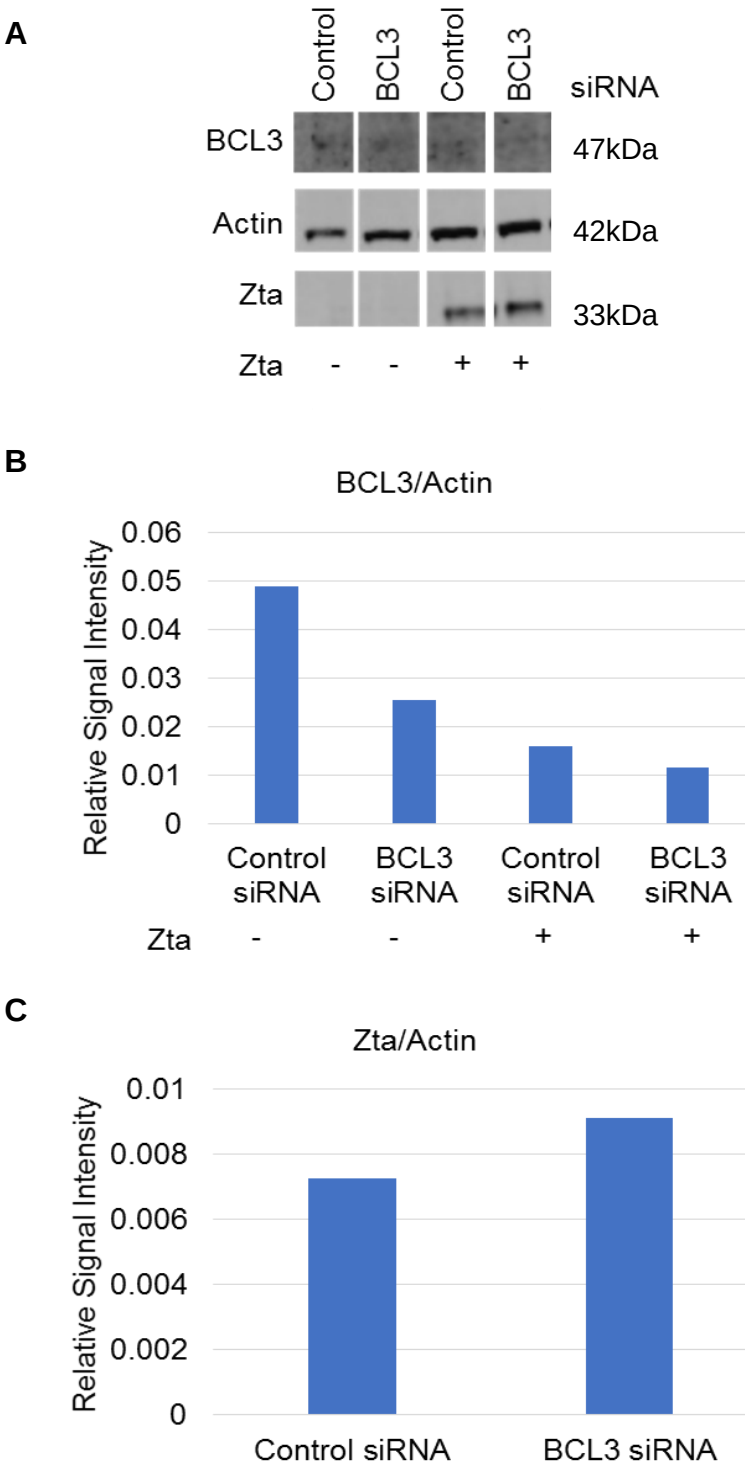


Figure 82: Western Blot (A) and corresponding band quantification for BCL3 (B) and Zta (C). HEK293ZKO cells transfected with a siRNA control or BCL3 specific siRNA followed by a Zta expression vector 24 hours later. Cells were harvested for western blotting and qPCR a further 48 hours later. Blot was stained using the Santa Cruz BCL3 (c-14) antibody.

It is not possible to observe EBV genome copy number increases 24 hours after induction of lytic cycle. The above experiment was repeated without ionomycin or PMA, as these led to substantial cell death. Cells were harvested 48 hours after the induction of EBV lytic cycle. The protein levels in the resulting samples were quantified with the help of a western blot and are shown in Figure 82. Results indicate that the BCL3 knock-down was successful and also confirm that less BCL3 is present in cells expressing Zta. The amount of Zta in cells transfected with BCL3 specific siRNA is 1.25 fold higher than in those transfected with a control siRNA.

The amount of EBV genome copy number in these cells was determined using qPCR (Figure 83). This shows a slightly higher amount of EBV in cells which received the BCL3 specific siRNA. However, this increase is not significant. It is possible that the increase in Zta abundance, detected when BCL3 is knocked down, correlates to an increase in EBV lytic cycle induction. 48 hours of induction may not be sufficient time to observe a significant effect on genome copy number. Harvesting cells 72 hours after transfection may make it possible to observe a more significant change.

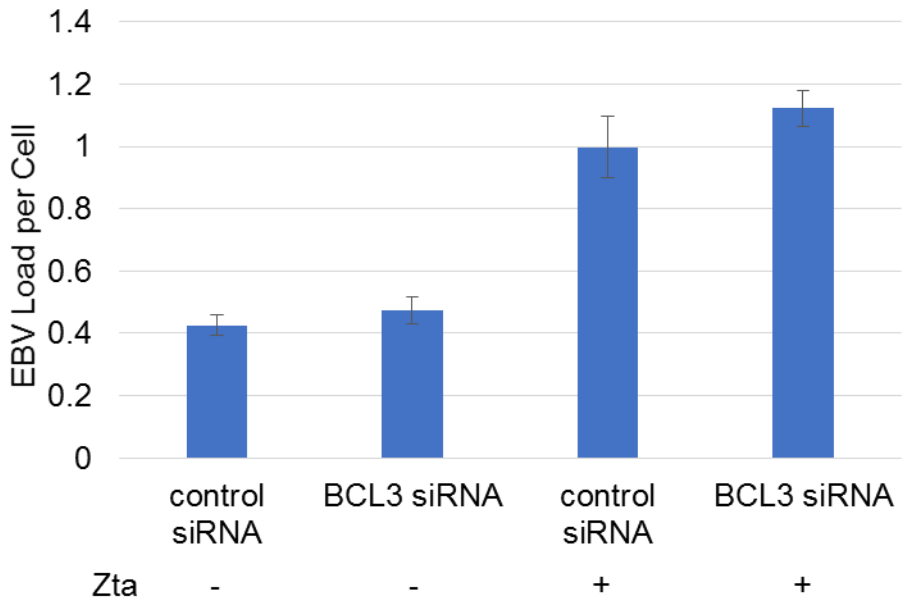


Figure 83: qPCR results showing the relative amount of EBV genome copies per human genome numbers. HEK293ZKO cells transfected with a siRNA control or BCL3 specific siRNA followed by the hisZta expression vector 24 hours later. Cells were harvested for western blotting and qPCR a further 48 hours later. Averages are based on triplicate pipetting repeats. Values are represented as mean ± SD.

5.6 BCL3 Over-Expression

To elucidate the effect of BCL3 on the reactivation of EBV from latency to lytic cycle by Zta, BCL3 was over-expressed in HEK293T cells with the help of a pcDNA3 plasmid expressing the 47kDa (454aa variant) isoform of the BCL3 protein.

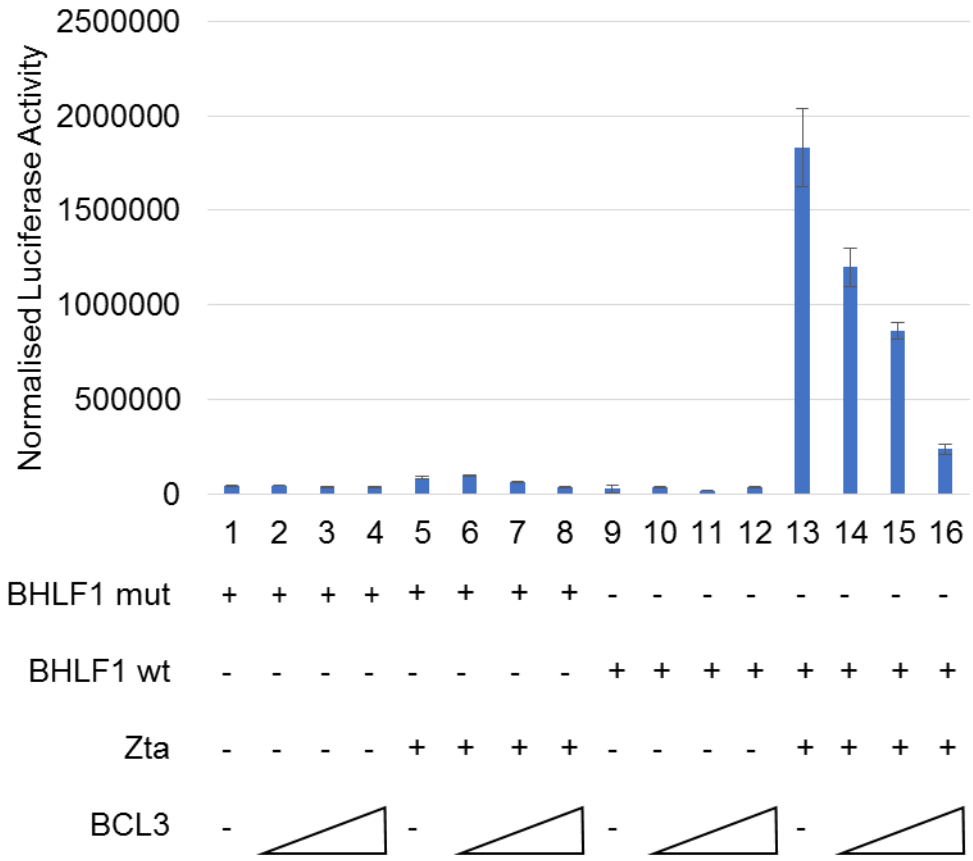


Figure 84: Luciferase assay results showing the expression of a BHLF1 or mutant reporter vector. The amount of luciferase intensity is proportional to the amount of Zta in each sample. HEK293T cells were transfected with increasing amounts of BCL3 flag (0.025µg, 0.1µg and 0.4µg), 0.1µg hisZta expression vectors and 0.5µg BHLF1 wild type or 0.5µg BHLF1 mutant Zta reporter vectors with a pcDNA3 vector to balance total transfect concentration to 2µg total. Cells were harvested 24hrs after transfection. The luciferase assay results were normalised to total protein concentration per sample as determined by BCA assay.

HEK293 T cells were transfected with a BHLF1 mutant or BHLF1 wild type reporter vector which creates luciferase intensity relative to the amount of Zta

activity. The BHLF1 reporter vector and mutant BHLF1 vector, containing mutated ZREs, which are described in more detail in [111].

In cells transfected with mutant BHLF1 (Figure 84, samples 1-8), only a low background of luciferase activity can be observed. In cells transfected with wild type BHLF1 and increasing amounts of BCL3, but without Zta (Figure 84, samples 9-12), a similar background luciferase intensity is shown. When wild type BHLF1 is transfected together with Zta, increased luciferase activity is observed while increasing amounts of BCL3 clearly decrease the amount of Zta activity (Figure 84, samples 13-16).

The western blot corresponding to the luciferase results is shown in Figure 85. The amount of Zta transfected into samples 5-8 and 13-16 was constant. However, as the amount of transfected BCL3 increases, the amount of Zta expressed appears to decrease, while overall cell number, as controlled for by Actin, remain constant.

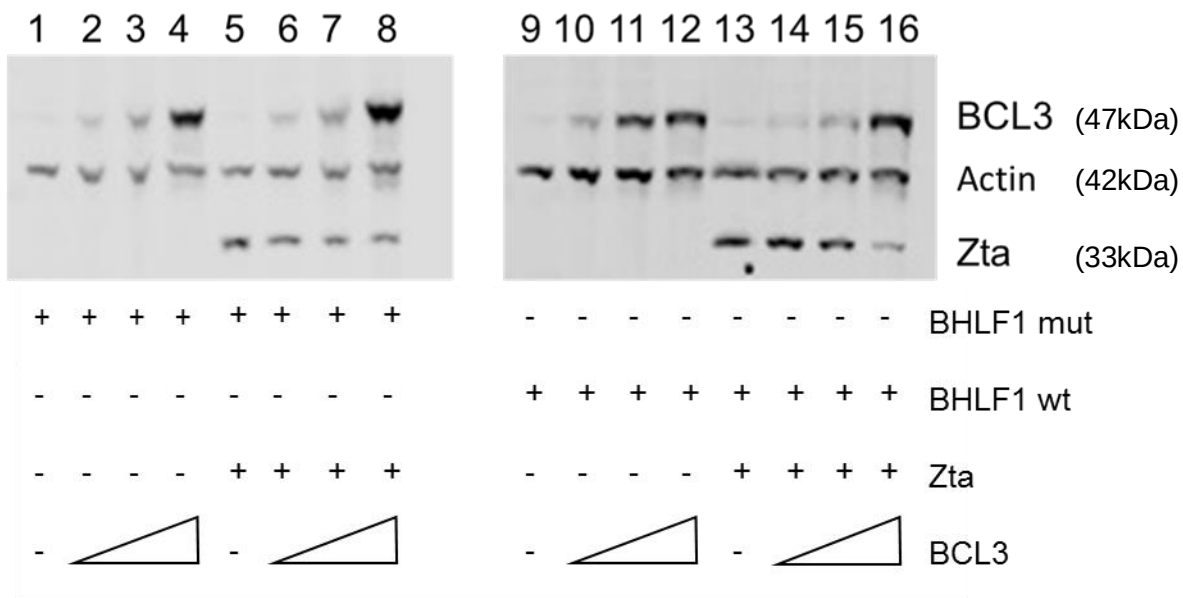


Figure 85: Western blots showing the expression of Zta and BCL3 with Actin as a loading control. This blot corresponds to the luciferase assay in Figure 84. HEK293T cells were transfected with increasing amounts of BCL3 flag (samples 1-4, 5-8, 9-12 and 13-16), 0.1µg hisZta expression vectors (samples 5-8 and 13-16) and 0.5µg BHLF1 wild type (9-16) or 0.5µg BHLF1 mutant (1-8) Zta reporter vectors with a pcDNA3 vector to balance total transfect concentration to 2µg total. Cells were harvested 24hrs after transfection. Blot was stained using the Santa Cruz (c-14) antibody.

The amounts of BCL3 and Zta were quantified and normalised relative to Actin levels. Figure 86 confirms that the amount of Zta protein decreases with increasing amounts of BCL3. The amounts of expression vector in each sample was normalised with a pcDNA3 vector.

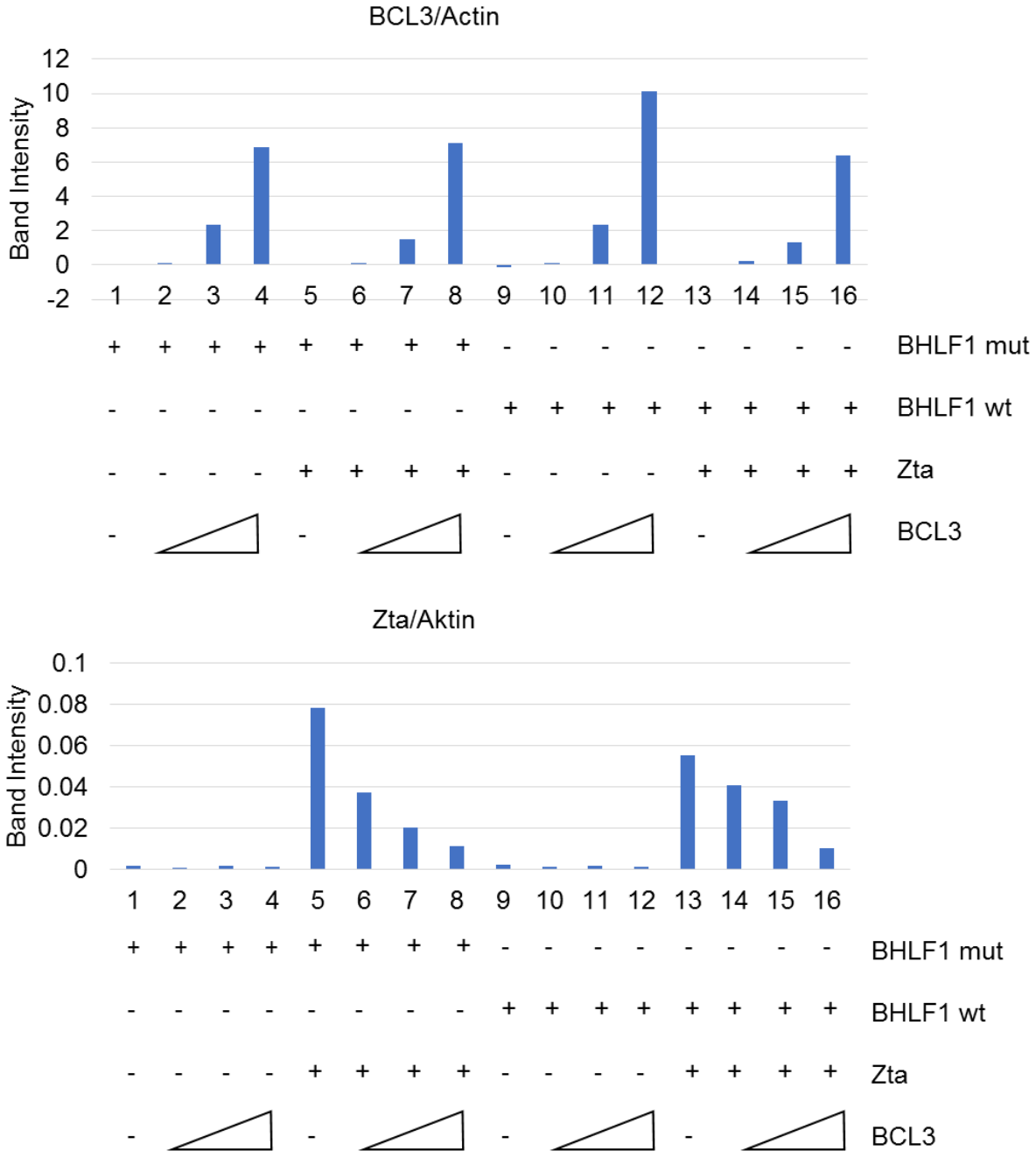


Figure 86: BCL3 and Zta western blot band quantification. Graphs show the amount of BCL3 and Zta expressed in HEK293T cells transfected with increasing amounts of BCL3 flag (samples 1-4, 5-8, 9-12 and 13-16), 0.1µg hisZta expression vectors (samples 5-8 and 13-16) and 0.5µg BHLF1 wild type (9-16) or 0.5µg BHLF1 mutant (1-8) Zta reporter vectors with a pcDNA3 vector to balance total transfect concentration to 1µg total. Expression levels are normalised to Actin.

To investigate the effect that BCL3 may have on lytic cycle reactivation, HEK293 ZKO cells were transfected with hisZta, BCL3, hisZta and BCL3, or a pcDNA3 control. The western blot shown in Figure 87 demonstrates that the transfections were successful. Consequently, the relative amount of EBV genome copy numbers was determined in the same lot of cells. This revealed a significant decrease in the amount of EBV in cells transfected with both, BCL3 and Zta, in comparison to cells transfected with Zta only (Figure 88). This drop in EBV genome copy number appears to be higher than the corresponding decrease of Zta abundance, suggesting that BCL3 is not the only factor which affects lytic cycle reactivation.

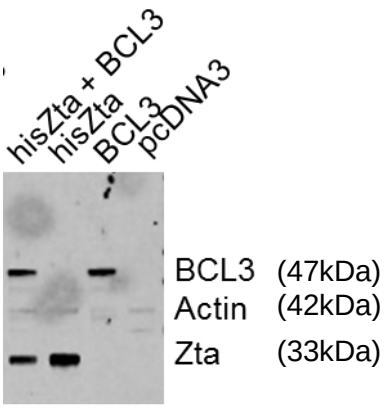


Figure 87: Western blot showing the expression of Zta and BCL3. HEK293ZKO cells were transfected with 100ng hisZta, BCL3, both or a pcDNA3 control vector and harvested 72hrs later. An antibody specific for actin was used as a loading control. Blot was stained using the Santa Cruz BCL3 (c-14) x antibody.

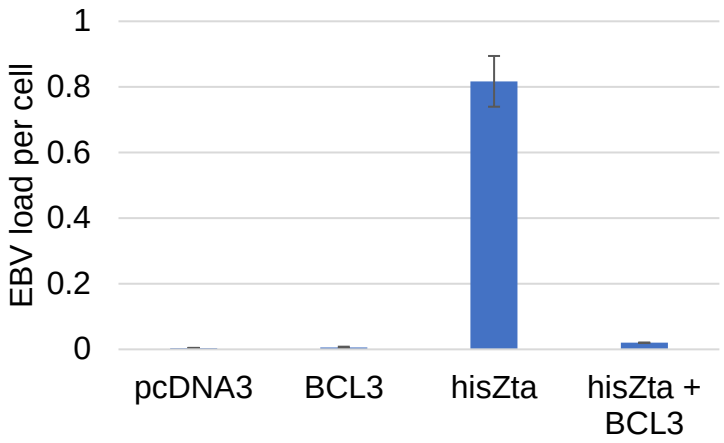


Figure 88: qPCR results showing the relative amount of EBV genome copies per human genome numbers. HEK293 ZKO cells were transfected with 100ng hisZta, BCL3, both or a pcDNA3 control vector and harvested 72hrs later. pcDNA3 was used to balance total vector amount to 200ng for each sample. The decrease in viral load per cell is significant when comparing transfection with Zta and BCL3 to Zta only ($p<0.05$). Averages are based on triplicate pipetting repeats. Values are represented as mean ± SD.

To determine what impact the BCL3 level has on the amount of lytic cycle inhibition, HEK293 ZKO cell were transfected with increasing amounts of BCL3 without (Figure 89, samples 1-4), or with constant amounts of Zta (Figure 89, samples 4-8). This revealed that the viral load per cell is significantly lower in cells transfected with both, Zta and large amounts of BCL3 (sample 8), in comparison to cells transfected with Zta only. The corresponding amount of Zta, BCL3 and Actin protein are shown in the western blot (Figure 90). The location of the BCL3 band corresponds to the expected size for the 47kDa isoform of the protein.

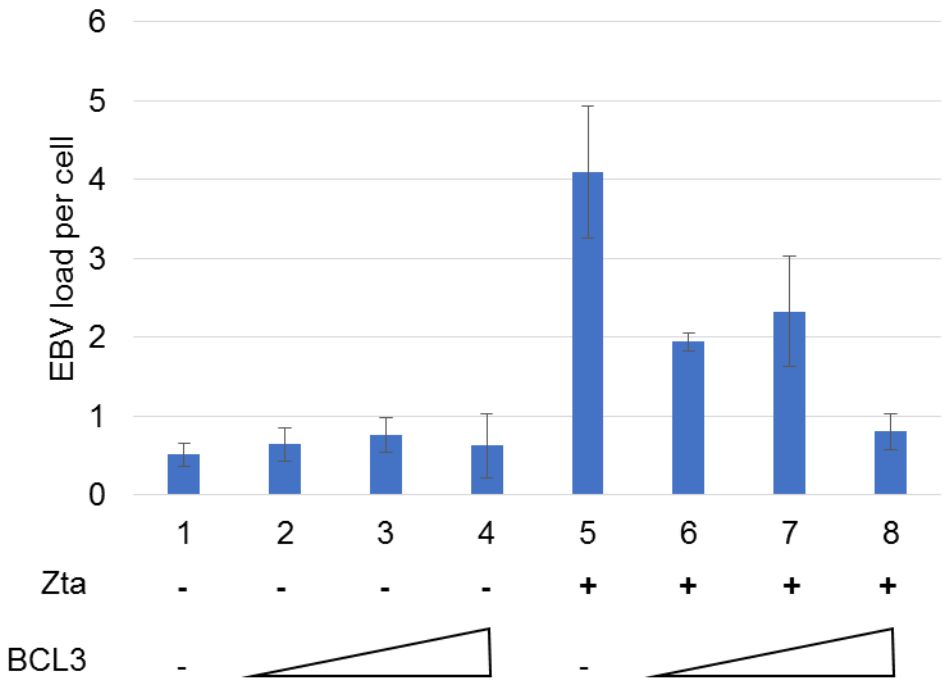


Figure 89: qPCR results showing the relative amount of EBV genome copies per human genome numbers. HEK293 ZKO cells were transfected with increasing amounts of BCL3 (samples 1-4 and 5-8) and 0.1µg hisZta expression vectors (samples 5-8) with a pcDNA3 vector to balance total transfect concentration. The viral load per cell is significantly lower in cells transfected with both Zta and large amounts of BCL3 (sample 8) in comparison to cells transfected with Zta only ($p \leq 0.05$). The decrease in EBV load is not significant in samples 6 and 7 when compared to sample 5. Cells were harvested 48hrs after transfection. Values are represented as mean ± SD.

The bands shown in the western blot (Figure 90) were quantified and the amount of Zta and BCL3 relative to actin levels can be seen in the graphs in Figure 91. This demonstrates that the amount of Zta decreases with increasing amounts of BCL3 even though the transfected amount of Zta is constant and the overall

amount of transfected vectors is equalised across each sample with the help of a pcDNA3 vector.

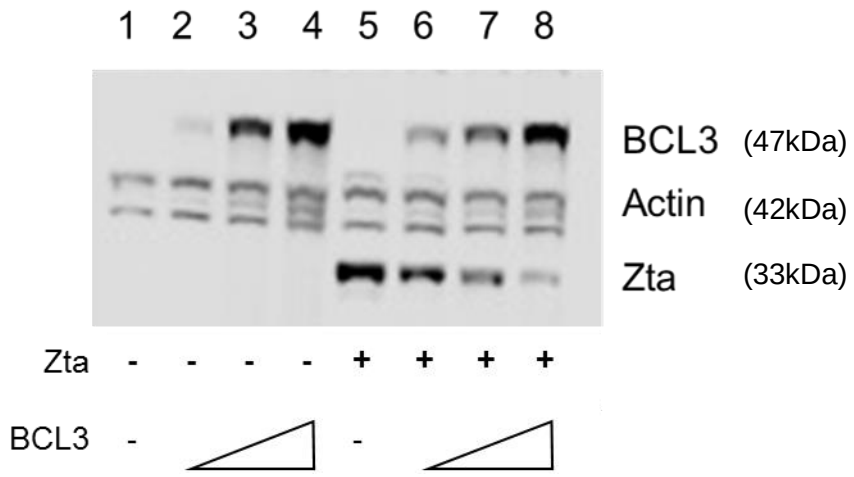


Figure 90: Western blot showing the expression of Zta and BCL3. HEK293 ZKO cells were transfected with increasing amounts of BCL3 flag (samples 1-4 and 5-8) and 0µg (samples 1-4) or 0.1µg hisZta (samples 5-8) expression vectors with a pcDNA3 vector to balance total transfect concentration to 2µg total. Cells were harvested 48hrs after transfection. An antibody specific for actin was used as a loading control. Blot was stained using the Santa Cruz BCL3 (c-14) antibody.

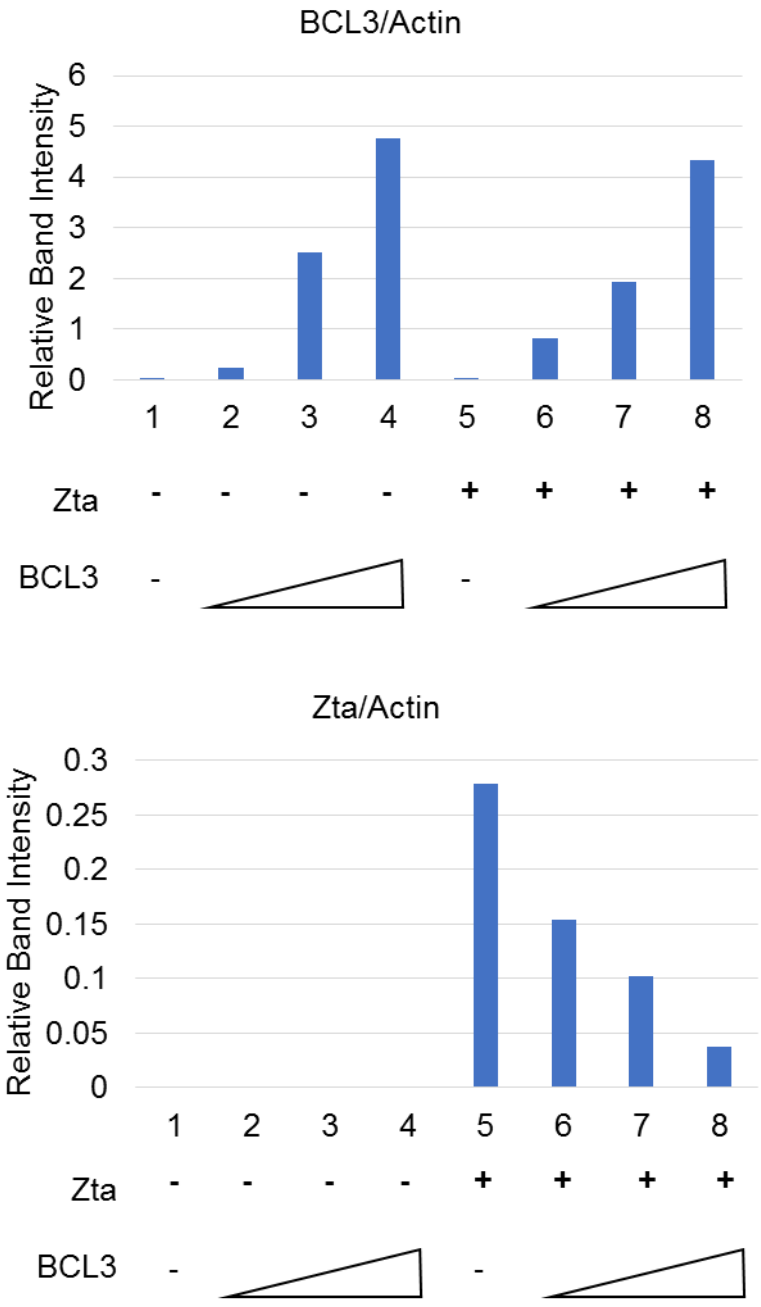


Figure 91: Western blot band quantification showing the amount of BCL3 and Zta. HEK293ZKO cells were transfected with increasing amounts of BCL3 flag (samples 1-4 and 5-8) and 0µg (samples 1-4) or 0.1µg hisZta (samples 5-8) expression vectors with a pcDNA3 vector to balance total transfect concentration to 1µg total. Cells were harvested 48hrs after transfection. Expression levels are normalised to Actin.

To determine whether BCL3 has an impact on lytic cycle other than decreasing the Zta protein level, a Zta titration experiment was performed in which HEK293 ZKO cells were transfected with a range of hisZta expression vector as well as a constant amount of Zta with BCL3 (Figure 92). This resulted in two samples which have very similar Zta protein levels, one transfected with Zta, whose expression is reduced by BCL3 (sample 6), and one transfected with a lower level of Zta to result in comparable protein levels (sample 3). The western blot showing BCL3, Zta and Actin protein levels in Figure 92 was quantified. The amount of Zta and BCL3 relative to Actin are shown in Figure 93.

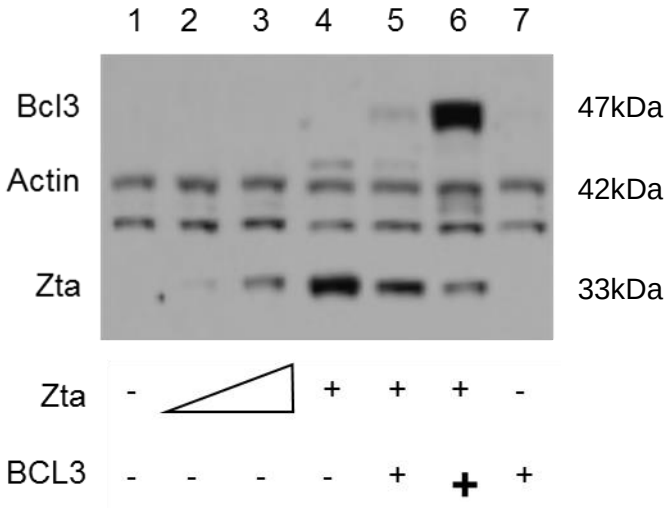


Figure 92: Western Blot showing the expression of Zta and BCL3. HEK293 ZKO cells were transfected with Zta (Samples 1 and 7: 0ng, samples 4-6: 280ng, sample 2: 30ng, sample 3: 60ng) and BCL3 (samples 1-4: 0ng, samples 5 and 7: 70ng, sample 6: 110ng). A pcDNA3 vector was used to balance total transfect concentrations across all samples. Cells were harvested 48hrs after transfection. An antibody specific for actin was used as a loading control. Blot was stained using the Abcam BCL3 antibody.

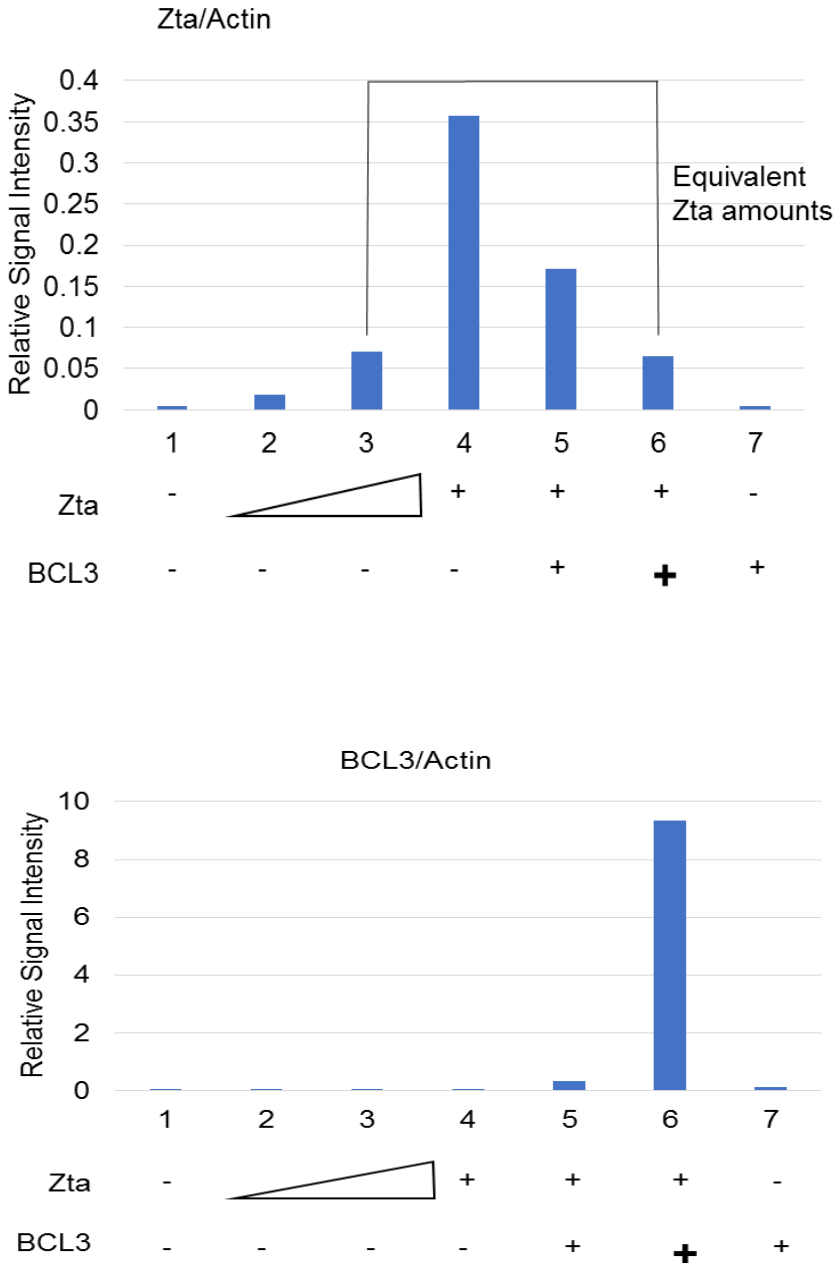


Figure 93: Western blot band quantification showing the amount of BCL3 and Zta. HEK293ZKO cells were transfected, as shown in the western blot above, with increasing amounts of hisZta (samples 1-4), or the highest amount of hisZta equivalent to sample 4 and increasing BCL3 flag (samples 5 and 6). Sample 7 was transfected with BCL3 flag only. A pcDNA3 vector was used to balance total transfect concentrations across all samples. Cells were harvested 48hrs after transfection. Zta levels were normalised to Actin levels.

The EBV genome copies per human genome were quantified corresponding to the protein levels in the previously described samples (Figure 94).

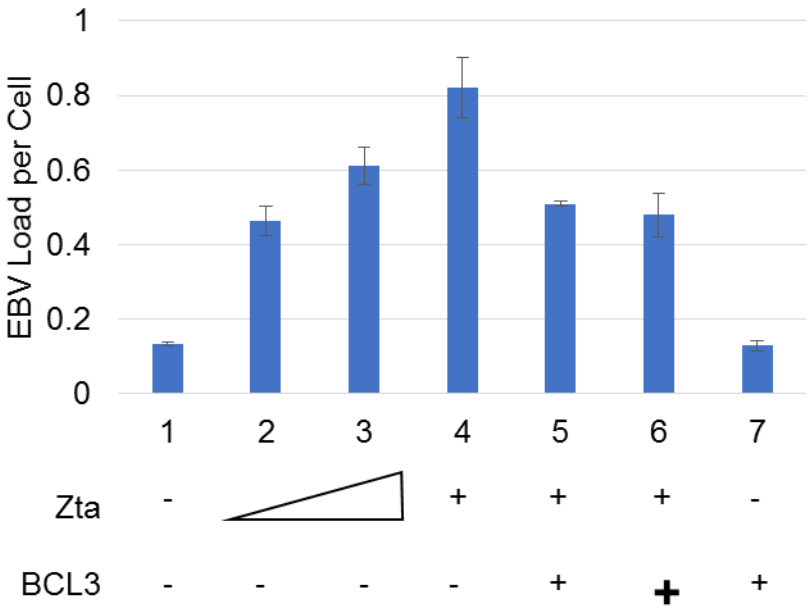


Figure 94: qPCR results showing the relative amount of EBV genome copies per human genome numbers. HEK293 ZKO cells were transfected with Zta (Samples 1 and 7: 0ng, samples 4-6: 280ng, sample 2: 30ng, sample 3: 60ng) and BCL3 (samples 1-4: 0ng, samples 5 and 7: 70ng, sample 6: 110ng). A pcDNA3 vector was used to balance total transfect concentrations across all samples. Cells were harvested 48hrs after transfection. In cells with the same expression level of Zta, viral load per cell is significantly higher when transfected with hisZta only (sample 3) in comparison to transfection with hisZta and BCL3 flag (sample 6, $p \leq 0.05$). Values are represented as mean ± SD.

The amount of Zta is comparable in samples 3, which was transfected with Zta only, and sample 6, whose Zta amounts were reduced to similar amounts by BCL3. Of these two samples, there is a small but significant difference in viral

load per cell when transfected with hisZta only in comparison to transfection with hisZta and BCL3. This result suggests that the attenuation of Zta function by BCL3 may relate to the decrease in Zta protein levels, but other factors are likely also involved in controlling EBV lytic cycle activation. This possibility was further explored in the following section.

5.7 Zta, BCL3, TORC and CBP

Data presented in the previous chapter showed that there is a high incidence of BCL3-Zta binding region overlap as well as a high incidence of H3K27me3 site overlap with Zta binding sites. To investigate whether these occurrences may be linked, the Zta binding sites common to Akata and HONE1-EBV cells were intersected with H3K27me3 sites in GM12878. A significant overlap of Zta binding sites which are also H3K27me3 sites and bound by BCL3 was discovered with almost 2/3 of the sites coinciding (Figure 95).

These results suggest that BCL3 may be acting as a repressor at 2/3 of sites which are bound by Zta. Therefore BCL3 may be involved in the establishment of a repressive chromatin environment, which is overturned or promoted by Zta.

Figure 95: Venn Diagram showing the number of H3K27me3 sites in GM12878 cells bound by Zta in HONE1-EBV and Akata cells overlapping with sites bound by BCL3 in GM12878 and Zta in HONE1-EBV and Akata cells. HONE1-EBV cells were induced with 10μM SAHA for 48 hours. Akata cells were induced with 0.125% IgG for 24 hours.

In order to further investigate the mechanism by which BCL3 may repress EBV lytic cycle activation, it was decided to investigate the possibility that one or more TORCs are involved in this process. It has been shown that both Zta and BCL3 interact with a member of the TORC family [90, 156]. In addition, the interaction of HTLV-1 Tax with BCL3 involved TORC and CBP during this virus' reactivation from latency [158] (see chapter introduction) and it is possible that the interaction of Zta with BCL3 occurs through a similar mechanism. CBP was one of the top five most significant factors which were shown to have common binding sites with Zta in epithelial and B cells. CBP, and its paralog p300, are transcriptional coactivators, central to eukaryotic transcription activation [14]. CBP provides a

scaffold for the interaction of a large variety of binding partners including >400 transcription factors and regulatory proteins [15], which compete for limiting quantities of CBP [16].

A series of western blots was performed to probe for the presence of TORC1-3. Unfortunately, it has not been possible to show BCL3, TORC1 or CBP expression by western blotting.

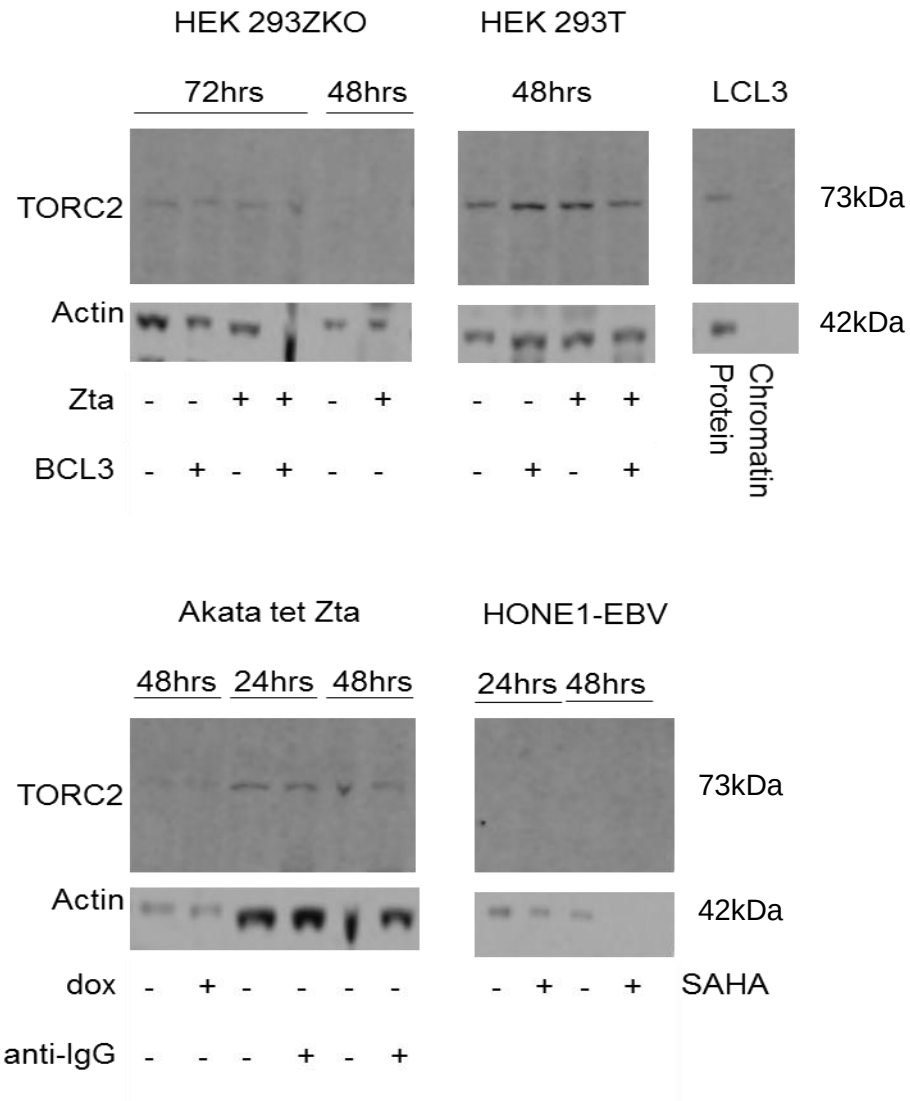


Figure 96: Western Blot showing the expression of TORC2. Figure shows a range of cell types and corresponding TORC2 expression (top row). Cells were treated to induce the EBV lytic cycle or Zta expression (HEK293T) and incubated for 24-72hrs as indicated. An antibody specific for actin was used as a loading control (bottom row).

The expression of TORC2 and TORC3, as well as Actin, in the indicated cell lines at various time points following induction are shown in Figure 96 and 97 respectively. TORC 3 could not be detected in Akata or HONE1-EBV cells (data not shown).

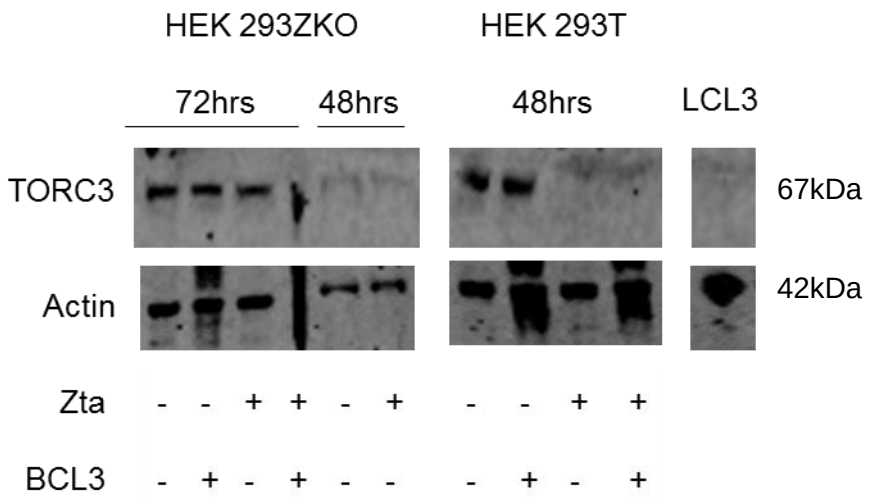


Figure 97: Western Blot showing the expression of TORC3. Figure shows a range of cell types and corresponding TORC3 expression (top row). Cells were treated to induce the EBV lytic cycle or Zta expression (HEK293T) and incubated for 48 or 72hrs as indicated. An antibody specific for actin was used as a loading control (bottom row).

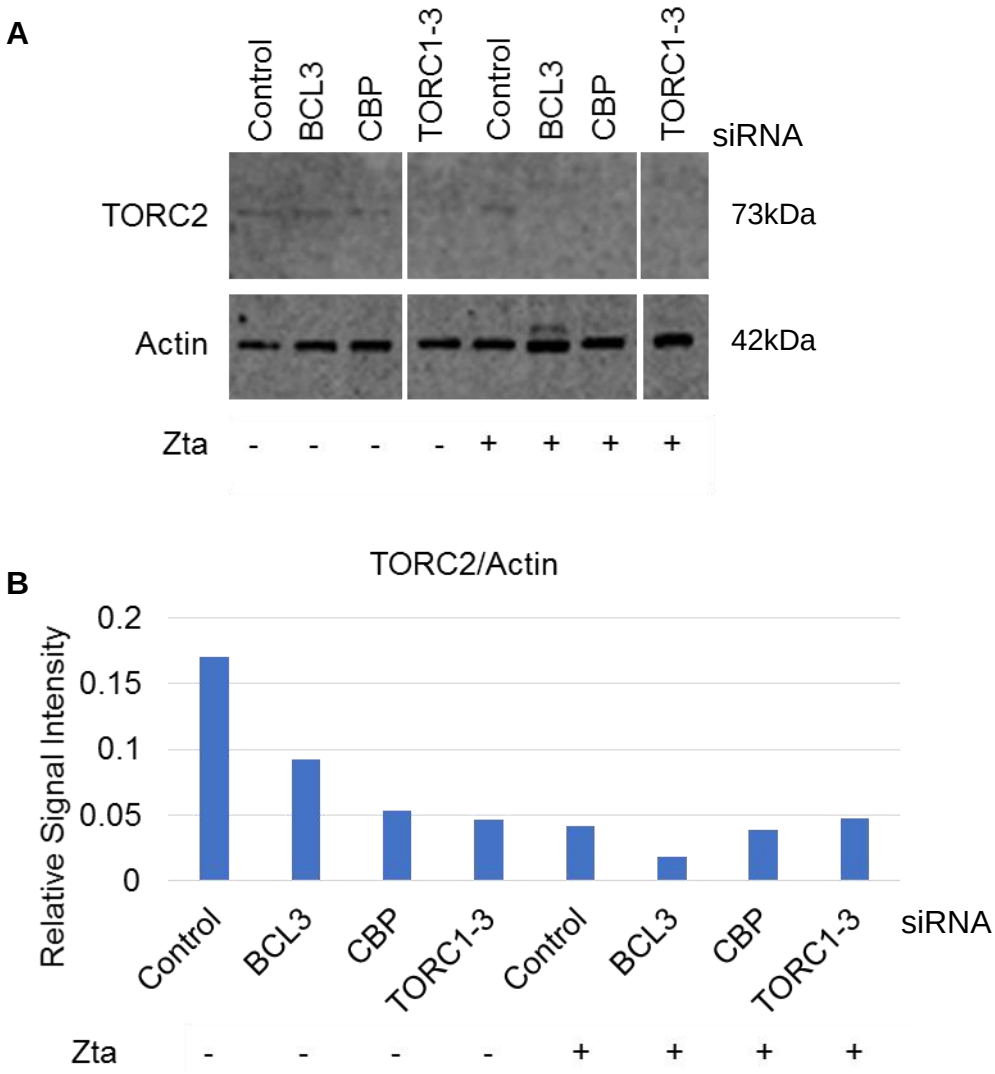


Figure 98: Western Blot (A) and corresponding band quantification (B) showing TORC2 and actin levels. HEK293 ZKO cells transfected with a siRNA control, BCL3, CBP or TORC1-3 specific siRNAs followed by the pcDNA3 or Zta expression vector and treatment with 1 µM ionomycin and 20ng/ml PMA 24 hours later. Cells were harvested for western blotting a further 24 hours later. The siRNA used in each sample is indicated. The first four samples were transfected with pcDNA3, the last four samples were transfected with Zta, as indicated.

TORC2 and 3 could be detected in HEK 293 cells and this cell type was chosen to perform siRNA transfections to investigate the effect of BCL3, CBP and TORC1-3 knock-down on Zta and EBV lytic cycle reactivation. Figure 98 suggests that transfection with BCL3 siRNA may reduce the abundance of TORC2 in HEK 293ZKO cells, which were treated with ionomycin and PMA to activate the B cell signalling pathway. The knock-down of TORC1-3 appears to have been successful in cells not transfected with Zta. A 3.7 fold reduction is shown (Figure 98 B). However, no decrease in TORC2/Actin can be observed when Zta is also present in the cells. Similarly CBP knock-down reduces TORC2/Actin abundance in cells not transfected with Zta. This reduction is not observed in the presence of Zta. However, a knock-down appears visible by eye. Due to the very low band intensity for TORC2 the signal:noise ratio may be too high to allow for dependable quantification of the relative signal intensity for TORC2/Actin.

The band intensity for TORC3 is higher than that for TORC2. As is shown in Figure 99, the knock-down of TORC3 was successful with a fold reduction of 2.4 in uninduced cells. Unlike TORC2, a decrease of TORC3/Actin is also observed in the presence of Zta, with a fold change of 3.3. Like TORC2/Actin, the abundance of TORC3/Actin is also reduced by CBP in the absence of Zta but not in its presence. Knock-down of BCL3 appears to reduce the level of TORC3/Actin to a greater extend in the presence of Zta.

To investigate the impact on EBV lytic cycle reactivation, the experiment was repeated without ionomycin and PMA, which leads to cell death, and cells were harvested 48 hours after transfected of the Zta expression vector. Western blot results for TORC2 are shown in Figure 100 and those for TORC3 in Figure 101.

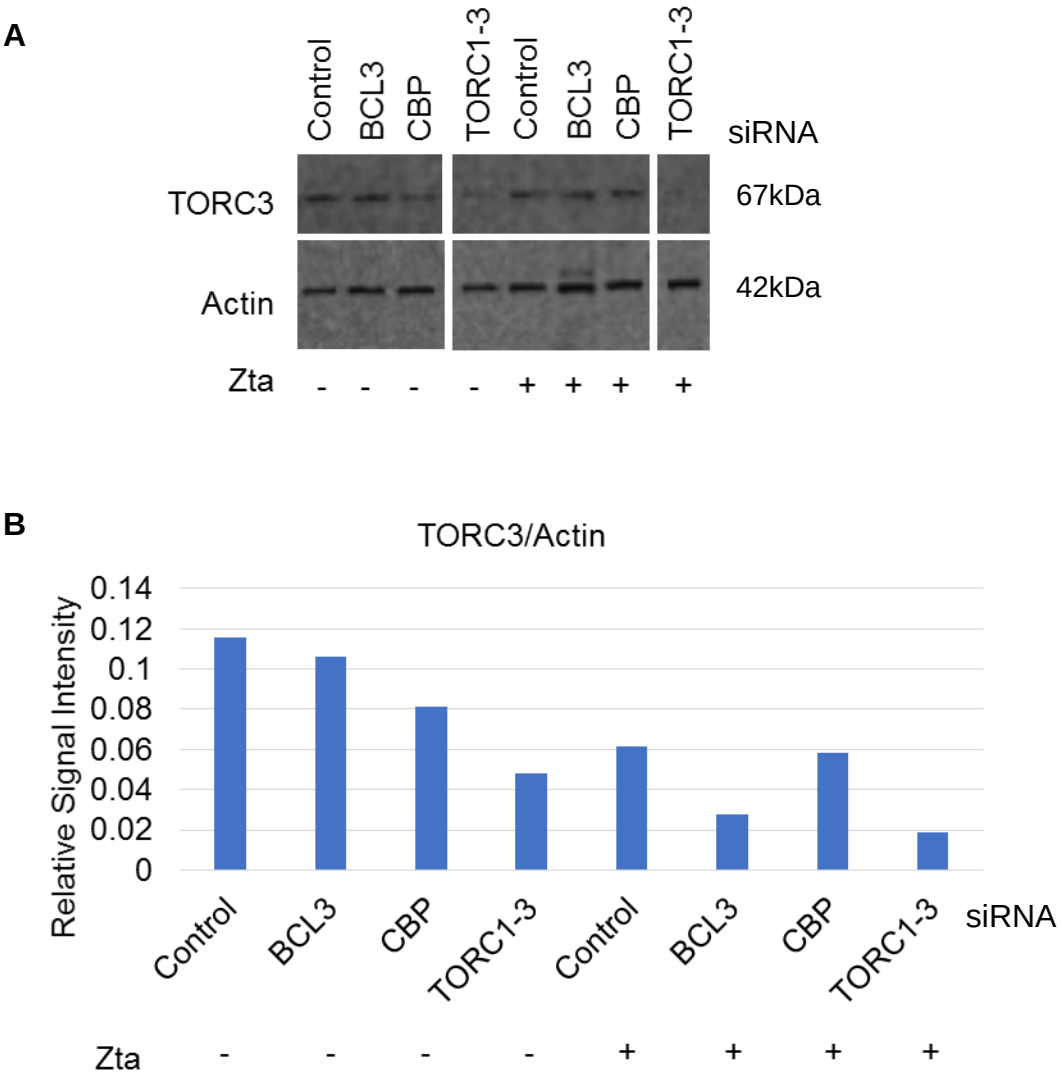


Figure 99: Western Blot (A) and corresponding band quantification (B) showing TORC3 and actin levels. HEK293ZKO cells transfected with a siRNA control, BCL3, CBP or TORC1-3 specific siRNAs followed by the pcDNA3 or hisZta expression vector and treatment with 1 µM ionomycin and 20ng/ml PMA 24 hours later. Cells were harvested for western blotting a further 24 hours later. The siRNA used in each sample is indicated above the blot. The first four samples were transfected with pcDNA3, the last four samples were transfected with Zta, as indicated.

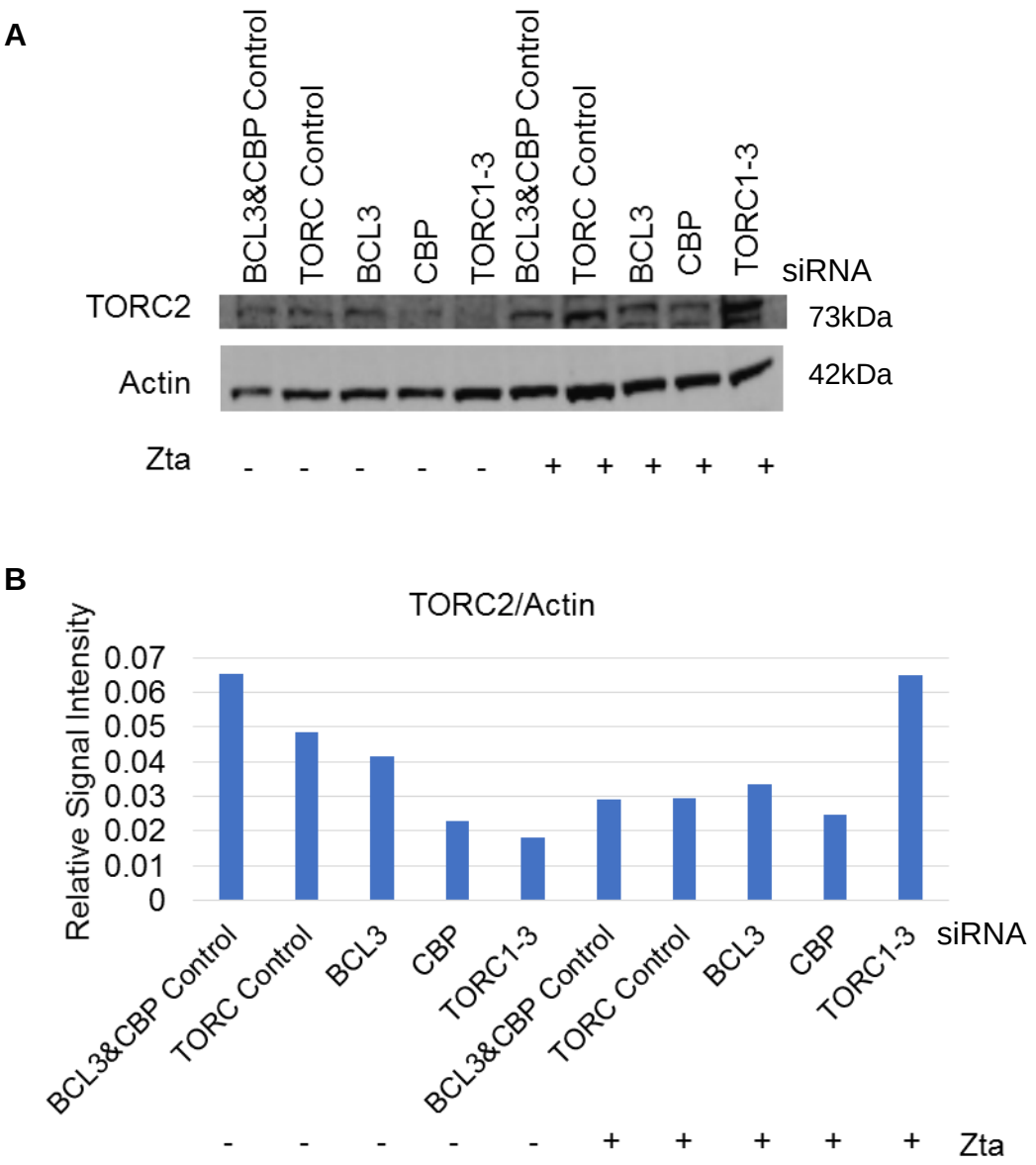


Figure 100: Western Blot (A) and corresponding band quantification (B) showing TORC2 and actin levels. HEK293ZKO cells transfected with a siRNA control, BCL3, CBP or TORC1-3 specific siRNAs followed by the pcDNA3 or hisZta expression vector 24 hours later. Cells were harvested for western blotting a further 48 hours later. The siRNA used in each sample is indicated above the blot. The first four samples were transfected with pcDNA3, the last four samples were transfected with Zta, as indicated.

The knock-down of TORC2 appears to have been successful in cells not transfected with Zta. However, the amount of TORC2 in cells containing Zta is much higher in cells transfected with TORC-specific siRNA in comparison to control siRNA. This experiment should be repeated.

Results for TORC3 demonstrate a knock-down of protein levels. When BCL3 is knocked-down, TORC3 levels also appear reduced when Zta is not present. It is possible that Zta stabilises TORC3 in a similar way as BCL3 may do, allowing TORC3 protein levels to stay constant in the absence of BCL3.

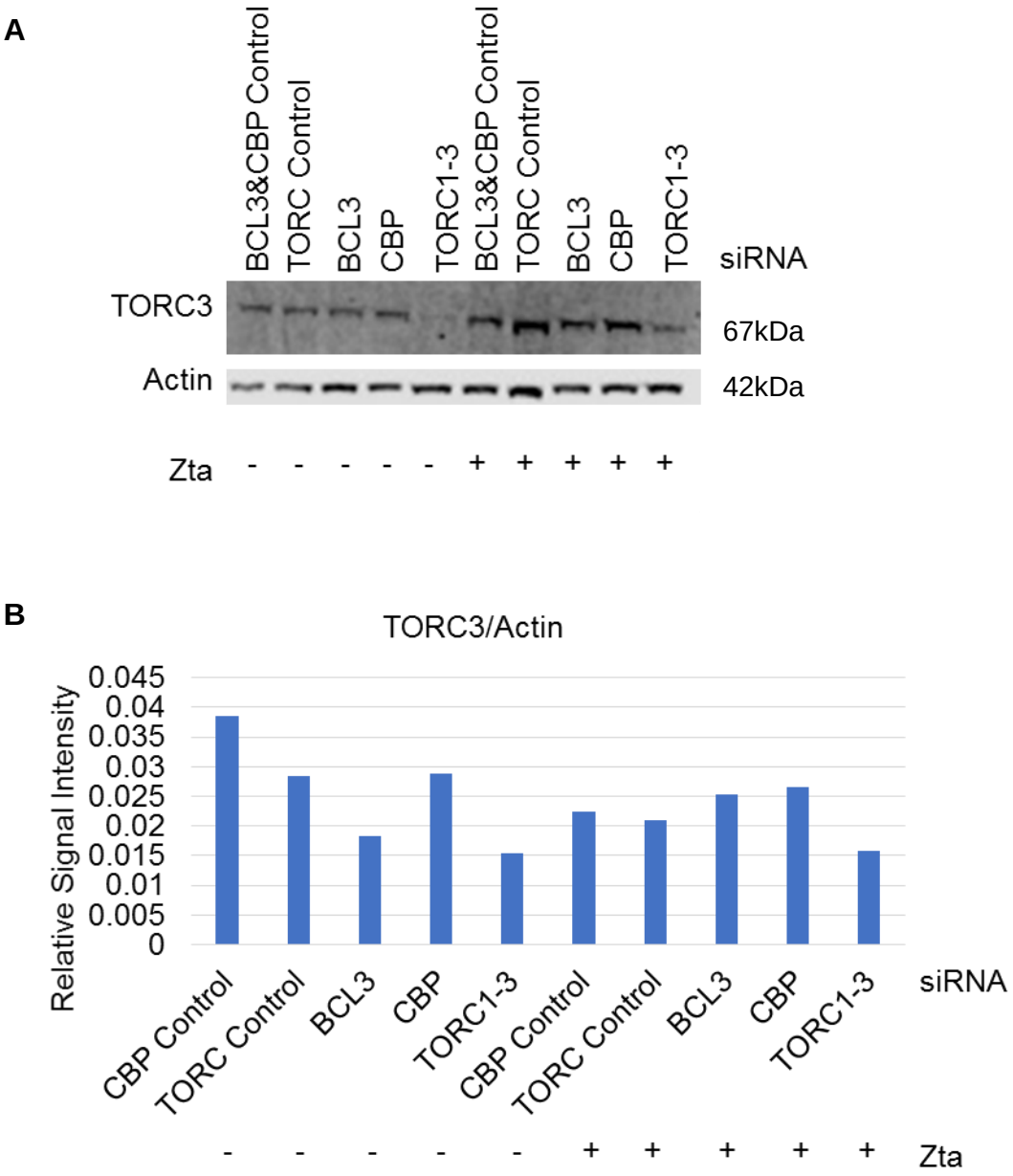


Figure 101: Western Blot (A) and corresponding band quantification (B) showing TORC3 and actin levels. HEK293 ZKO cells transfected with a siRNA control, BCL3, CBP or TORC1-3 specific siRNAs followed by the pcDNA3 or hisZta expression vector 24 hours later. Cells were harvested for western blotting a further 48 hours later. The siRNA used in each sample is indicated above the blot. The first four samples were transfected with pcDNA3, the last four samples were transfected with Zta, as indicated.

To investigate whether the knock-down of TORC1-3 and CBP has any effect on Zta abundance or EBV reactivation from lytic cycle, a western blot was performed with cells harvested 24 hours after transfection of Zta (Figure 102 A) and the amount of Zta was normalised to actin levels (Figure 102 B). This showed a very slight reduction of Zta. When this experiment was repeated without ionomycin or PMA and protein levels were investigated 48 hours after transfection with a Zta expression plasmid, the reduction in Zta protein level as a result of CBP and TORC knock-down is much clearer (Figure 103).

Figure 104 shows the impact on lytic cycle activation resulting from the transfection of CBP or TORC1-3 specific siRNAs. Knock-down of either CBP or TORC1-3 leads to a significant decrease of EBV lytic cycle activation 48 hours after transfection with Zta.

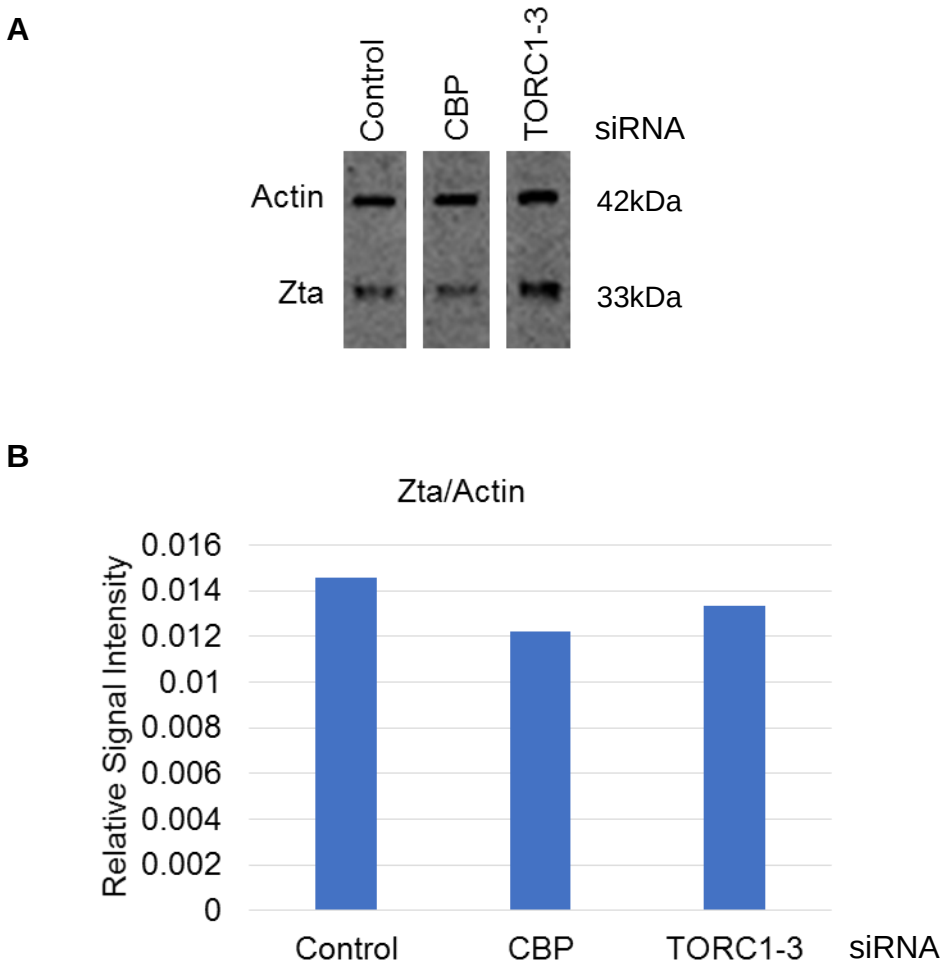


Figure 102: Western Blot (A) and corresponding band quantification (B) showing Zta and actin levels. HEK293 ZKO cells transfected with a siRNA control, CBP or TORC1-3 specific siRNAs followed by Zta expression vector and treatment with 1 µM ionomycin and 20ng/ml PMA 24 hours later. Cells were harvested for western blotting a further 24 hours later. The siRNA used in each sample is indicated.

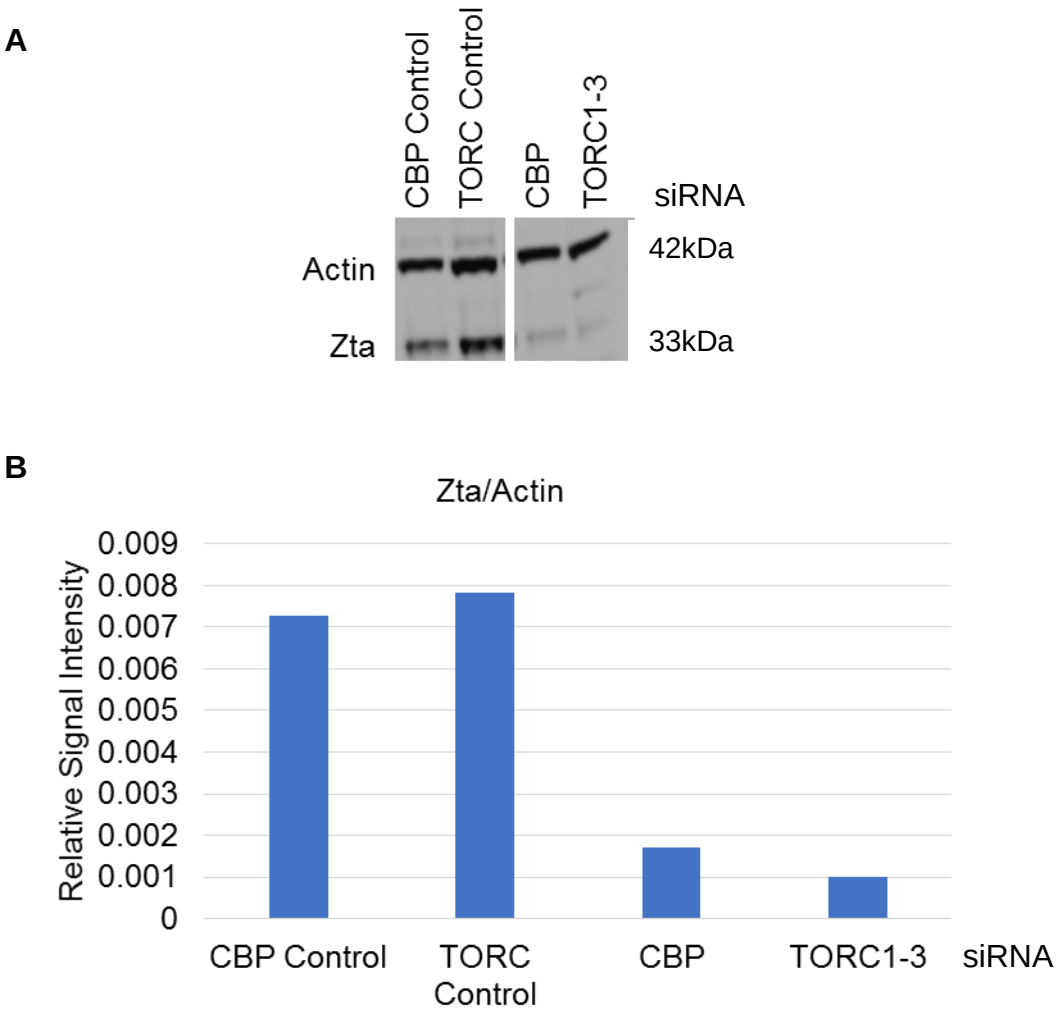


Figure 103: Western Blot (A) and corresponding band quantification (B) showing Zta and actin levels. HEK293 ZKO cells transfected with a siRNA control, CBP or TORC1-3 specific siRNAs followed by hisZta expression vector 24 hours later. Cells were harvested for western blotting a further 48 hours later. The siRNA used in each sample is indicated.

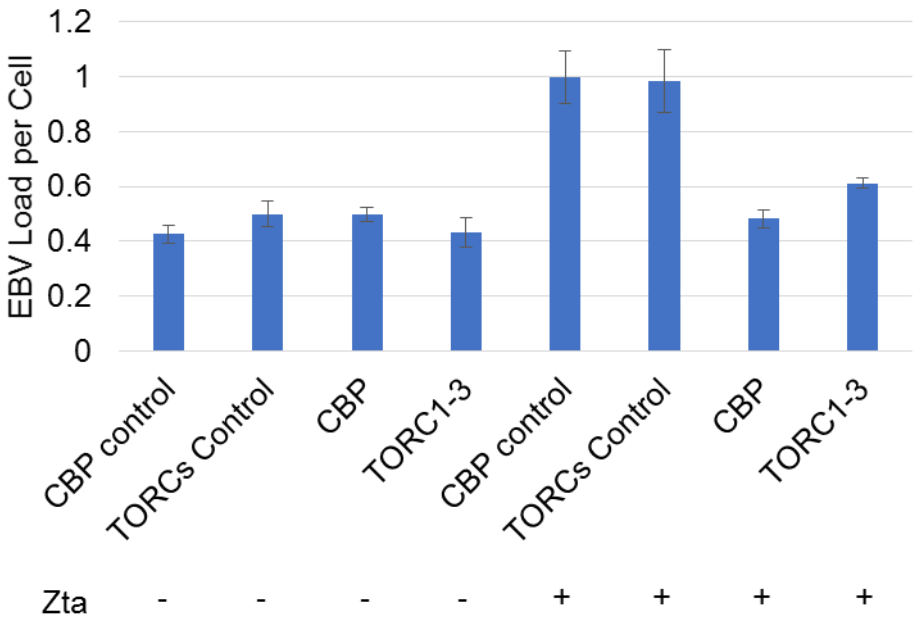


Figure 104: qPCR results showing the relative amount of EBV genome copies per human genome numbers. HEK293ZKO cells were transfected with a siRNA control, CBP or TORC1-3 specific siRNAs followed by hisZta expression vector 24 hours later. Cells were harvested for western blotting a further 24 hours later. The siRNA used in each sample is indicated above the blot. The first four samples were transfected with pcDNA3, the last four samples were transfected with hisZta, as indicated. The difference in EBV load per cell is significantly smaller in cells transfected with CBP or TORC1-3 siRNA in comparison to transfection with the relevant control siRNA ($p<0.05$). Averages are based on triplicate pipetting repeats. Values are represented as mean ± SD.

5.8 Discussion

This chapter summarises results which suggest that BCL3 plays a part in the EBV lytic cycle.

It was demonstrated that BCL3 is expressed in the cell lines which Zta ChIP sequencing results were obtained in. In addition, an increase in BCL3 protein level was observed as a result of lytic cycle induction in a range of cell lines. No corresponding increase in BCL3 RNA levels could be observed, suggesting that post-transcriptional processing is responsible for the increase in BCL3 protein levels. On the other hand, the RNA level of p50 is significantly increased in the epithelial cell line HONE1-EBV but not in Akata cells when EBV is in lytic cycle. The p50/BCL3 heterodimer has been shown to be the major form for NFκB detected in Nasopharyngeal Carcinoma [148]. This suggests that there may be an epithelial specific mechanism during lytic cycle involving this NFκB subunits.

ChIP sequencing data analysis discussed in the previous chapter suggested that BCL3 binds to many of the regions of the human genome that Zta also binds to. This was confirmed for several binding sites with the help of ChIP qPCR.

Further ChIP experiments in LCL3, followed by a second round of ChIP and qPCR, did not show BCL3 and Zta binding to these common binding sites at the same time.

In addition to investigating BCL3 and Zta on the human genome, BCL3 ChIP sequencing data mapped to the viral genome by Arvey *et al.* was investigated [171]. It was not possible to show clear evidence of BCL3 binding to the viral genome in Akata cells because of a lack of negative control flanking region. As there is no evidence for Zta and BCL3 binding at the same time, it is possible that

Zta displaces BCL3. Therefore BCL3 may not commonly bind to the viral genome during lytic cycle.

It has proven challenging to quantify the increase in BCL3 protein levels observed in Akata cells as a result of lytic cycle induction using FACS. This is partly due to the lack of a suitable non-rabbit BCL3 antibody which would allow detection of BCL3 without interference of the rabbit IgG antibody used to induce lytic cycle in Akata cells. A different cell line with a different mode of EBV lytic cycle induction may also be an alternative in future experiments.

In order to investigate the effect of BCL3 on the EBV lytic cycle, BCL3 specific siRNA was used to attempt to knock-down the expression of BCL3. Initial optimisation of the experiment showed promising transfection results, as shown via FACS with a fluorescent Cy3 marker siRNA 24 hours after transfection. However, it was not possible to show a knock-down of BCL3 RNA, possibly due to a degradation of the siRNA over time.

Using a new siRNA and transfection system to knock-down BCL3 protein levels coincided with an increase in Zta abundance at 24 and 48 hours after transfection of HEK293 ZKO cells with a Zta expression vector. In addition, a slight increase in EBV genome copy numbers could be detected. While this increase is not significant, based on the increase in Zta levels, it is possible that it would be at a later time point when lytic cycle had progressed further as the effect of BCL3 over-expression was shown to be significantly larger at 72 hours than 48 hours after transfection. However, it was not possible to clearly show that transfection with BCL3 specific siRNA lead to a knock-down of BCL3. More experiments are required to demonstrate this.

BCL3 was over-expressed in HEK293T and HEK293 ZKO cells. This showed a progressive decrease in Zta protein levels as a result of increasing BCL3 amounts. The amount of EBV lytic cycle is decreased as a result of BCL3 transfection. This effect is more prominent at 72 hours after transfection than at 48 hours after transfection. The decreased abundance of Zta when BCL3 is over-expressed may relate to the attenuation of EBV lytic cycle. However, Zta transfected without BCL3 results in significantly higher EBV genome copy numbers per cells than Zta co-transfected with BCL3 to obtain similar Zta protein levels. Furthermore, higher amounts of BCL3 do not seem directly proportional to lower amounts of lytic cycle.

As both, Zta and BCL3 have been shown to interact with TORC, the possibility that TORC is involved in the mechanism by which BCL3 inhibits the lytic cycle was investigated by siRNA knock-down of TORC1-3. CBP was also included in this investigation as it is known that CBP is targeted by carcinogenic viruses to hijack the cellular regulatory machinery, which disrupts host signalling pathways [17]. In addition, CBP was shown to be one of the top five co-associations with Zta binding peaks common to HONE1-EBV and Akata cells (Chapter 4).

BCL3 knock-down decreases TORC protein levels at 24 and 48 hours after Zta or pcDNA3 expression vector transfection, unless Zta is also present at 48 hours. The decrease of TORC2 and 3 can be seen in the presence of Zta at 24 hours, but not anymore by 48 hours. It is possible that both Zta and BCL3 stabilize TORC proteins.

The levels of TORC2 do not appear to decrease when cells are transfected with TORC2 specific siRNA when Zta is also expressed. More experiments are necessary to confirm if this result is reproducible.

The impact of TORC and CBP knock-down on Zta protein levels is much higher at 48 hours after Zta transfection than at 24 hours, suggesting that these proteins may play an important role in the establishment of lytic cycle. Close attention to the timing of further investigations into this process will be essential.

A model of how Zta interacts with TORC and CBP, and how BCL3 may interfere with this mechanism is shown in Figure 105. Zta has been shown to interact with CBP as well as TORC2 [90, 98]. CBP can also bind to cellular transcription factors including p65, p53 and the retinoic acid receptor [84, 85]. p65, p53 and the retinoic acid receptor inhibit Zta transcriptional activity, and the over-expression of CBP relieves the antagonistic interactions between different classes of activators such as nuclear receptors and AP1. It has been suggested that CBP availability may control the transcription output from a range of signal inputs [84].

The mechanism by which BCL3 inhibits the lytic cycle of EBV may be similar to that described for HTLV-1 in the introduction of this chapter, which involves blocking the binding of CBP to the complex, thus decreasing transcription activation. Alternatively CBP may be present, but the binding of BCL3 to the complex may block the activation of gene transcription (Figure 105 C). As TORC is a tetramer, it is also possible that multiple BCL3 and Zta proteins could bind to the complex to either inhibit or promote its transcription activation function. However, if this is the case it would be expected that the level of BCL3 is more proportional to the level of attenuation, which has not been shown to be the case.

These versions of the model would involve BCL3 binding to Zta binding sites at the same time as Zta. ChIP reChIP experiments described in this thesis did not show any evidence of this. In addition, the ChIP sequencing data which shows BCL3 at Zta binding sites on the human and viral genome was obtained in cells which are reported to contain latent EBV infection.

It may be the case that BCL3 associates with Zta binding sites during latency, possibly as part of a complex including TORC and CBP. BCL3 may recruit HDACs to this complex. This model (Figure 105 C) is supported by a large overlap of Zta binding sites, which are also H3K27me3 sites and bound by BCL3 with almost 2/3 of the sites coinciding. It is possible that Zta displaces BCL3 at such a site, thus activating gene transcription and the lytic cycle. As BCL3 does not bind DNA directly, this model would involve the presence of an additional factor (X) at Zta binding sites during latency. This factor may be CREB, p50 or p52 as BCL3 has been shown to interact with these proteins during its role as transcription repressor [156]. AP1 or a viral factor could also tether BCL3 to DNA at Zta binding sites. The fact that indirect BCL3 binding motif analysis in a EBV negative cell line (K562) did not reveal a non CpG ZRE, while the indirect BCL3 binding motif in GM12878 did resemble a non CpG ZRE may point to a latent viral factor being involved in this process. Further research would be required to specify these details.

BCL3 ChIP qPCR experiments in HEK293 ZKO cells, which are tightly latent unless transfected with Zta, may be able to reveal if BCL3 is present at certain binding sites, like OriLyt, in latency or in lytic cycle, or both.

BCL3 has been shown to interact with TORC3 via its ankyrin repeat region. Interestingly, Dryfus *et all*. suggest that Zta may also have an ankyrin-like region [155], so it would be possible that Zta may interact with TORC in this way instead of through the bZIP domain. This could be an alternative way of explaining how BCL3 is inhibiting Zta and lytic cycle activation by competing for limiting amounts of TORC outside of a complex bound to DNA. However, this model would not explain the presence of BCL3 at Zta binding sites.

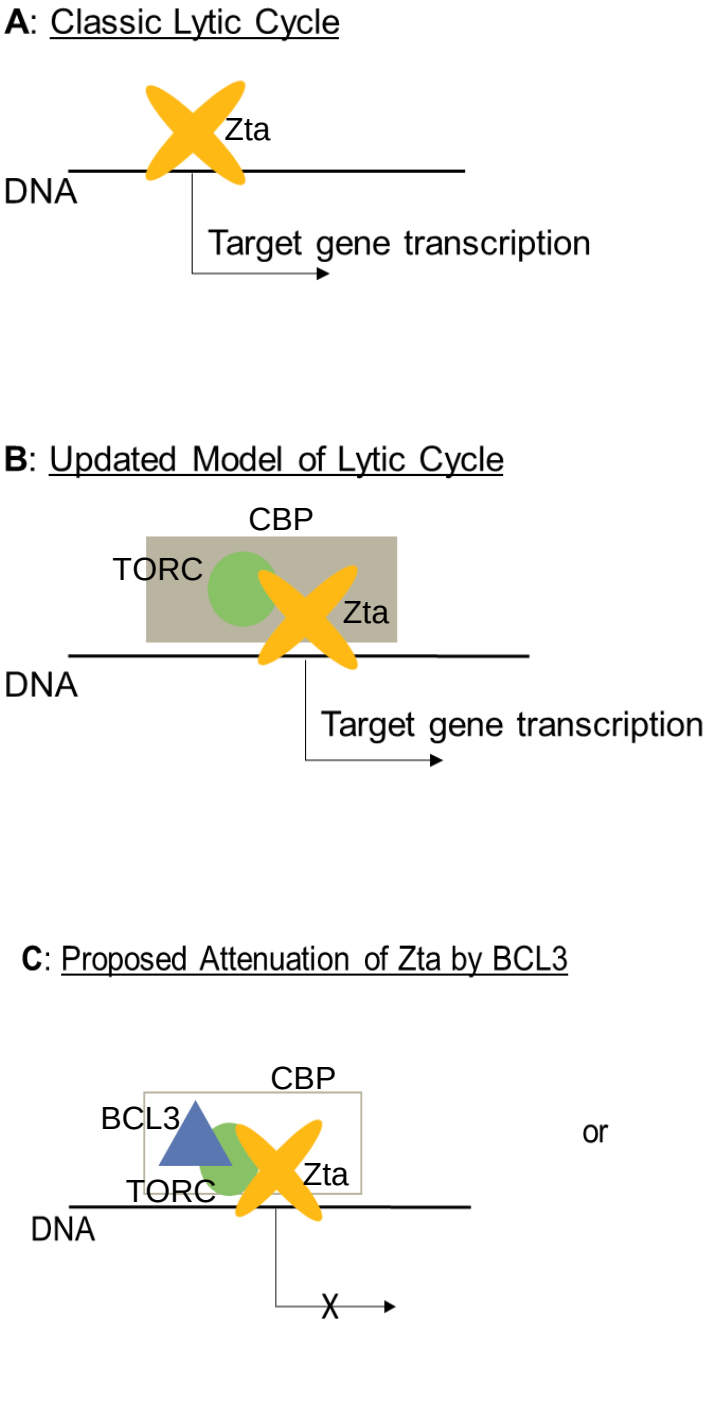


Figure 105: Model of Zta, CBP and TORC interaction and how BCL3 may interfere with it. A: Model of classic target gene activation by Zta. B: Proposed model of Zta gene activation. C: Different possible mechanisms by which BCL3 may block Zta transcription activation.

Chapter 6: General Discussion

Since the discovery of EBV in 1964, our understanding of the virus has evolved from being a contributing factor in an obscure African cancer to its present leading role as a prime example of a human tumour virus. EBV infects a large majority of the human population and is linked to a very diverse range of diseases. Considering the significant global burden of EBV-associated malignancies, it is an important priority to find novel treatments with specific targets that exploit the presence of the virus in malignant cells [2]. Although the precise role of EBV in epithelial malignancies is still poorly understood, the fact that the virus is present in all tumour cells shows that there is opportunity for the development of novel therapeutic and diagnostic approaches [6].

The EBV early lytic cycle protein Zta is a key regulator of the switch from latency to lytic cycle. This thesis presents new insights into the binding, and consequent control, of Zta on the host cell genome. For this purpose, a ChIP sequencing experiment was performed in epithelial cells. The challenges encountered during the large scale Zta ChIP and preparation of a library for sequencing enabled the writing of a comprehensive book chapter on the methods involved in ChIP sequencing for the analysis of Zta binding to the human and EBV genome (Appendix B) [172].

Sequencing reads mapped to the human and EBV genome were analysed to obtain MACS peaks, binding motifs, gene associations and gene list enrichments. In addition, results were compared to those previously obtained in B cells. This revealed Zta binding sites which could be cell-type specific, and some that could be common. Further research is needed to ascertain the relevance of the common and unique binding peaks of each cell type. However, as the rapid

advance in sequencing technology resulted in epithelial cell ChIP sequencing data to be of higher quality in comparison to the results obtained in B cells several years earlier, it is likely that at least part of the unique peaks found in epithelial cells are lower significance peaks also present, but not identified, in B cells.

Histone modification and transcription factor binding data, which is publicly available from the ENCODE database for binding to the human genome, were included in the Zta ChIP sequencing data analysis. This revealed a range of co-associations, of which BCL3 had the notably highest significance. The intersection of Zta and BCL3 MACS peaks in GM12878 cells, which are reported to contain latent EBV infection, confirmed a high number of overlaps. Interestingly, this overlap is especially high for Zta peaks common to B and epithelial cells, which could point to a cell-type independent role of BCL3.

In addition to ChIP sequencing data mapped to the human genome, BCL3 binding mapped to the EBV genome (Arvey *et al.* [171]) showed a BCL3 binding peak at OriLyt, which is also known to be bound by Zta. BCL3 ChIP qPCR results in Akata cells, containing lytic EBV, showed enrichment of OriLyt by a BCL3-specific antibody in comparison to an IgG control. However, while amplification above the IgG negative control can be seen for BCL3 in Akata cells, no negative control flanking region is available for BCL3 binding to this peak. Therefore, amplification could represent a false positive caused by non-specific binding of the BCL3 antibody used to perform the ChIP. Further experiments are necessary to validate this binding.

In HONE1-EBV cells induced with SAHA, and Akata cells induced with anti-IgG, it was shown that BCL3 protein levels increase when EBV is in lytic cycle and Zta

is expressed. It is not clear if this effect is due to the direct impact of Zta on BCL3 stability and/or translation, or due to the impact of SAHA and anti-IgG on both, Zta and BCL3 expression. On the other hand, results obtained with the help of luciferase assays in HEK 293T cells show that Zta transcription activation of a BHLF1 reporter vector is decreased when BCL3 is over-expressed. Furthermore, EBV genome copy number qPCRs performed in HEK 293ZKO cells showed that EBV lytic cycle activation is also decreased by BCL3 over-expression. Both of these consequences of BCL3 over-expression correlate with a decrease in Zta protein levels. This may be caused by BCL3 impacting on the expression and/or stability of Zta. However, further experiments are necessary to investigate this correlation. Zta and BCL3 were both transfected and expressed in HEK 293 cells with the help of pcDNA3 Zta and pcDNA3 BCL3 expression vectors. It is possible that native Zta expression is not affected in the same way. Future work should question the impact of BCL3 on Zta expression and stability.

As outlined in the introduction for Zta and BCL3, both proteins have been shown to interact with a member of the TORC family. In addition, BCL3 and TORC are involved in the reactivation from lytic cycle of the human T cell leukaemia virus HTLV-1. HTLV-1 is a lymphotropic virus, like EBV, and also causes life long, latent infections. The HTLV-1 protein Tax acts as a switch from latency to lytic cycle. This process involves the association of Tax with CREB. It is not clear to what degree TORC1, TORC2 and TORC3's functions are similar or interchangeable. While they are not homologs, all three members of the TORC family significantly enhance Tax-mediated transcription by forming a complex with Tax and CREB. Low-level expression of each TORC is present in all human tissues, however, TORC2 and TORC3 are predominantly found in T and B

lymphocytes [3]. The interaction of BCL3 with TORC3 significantly inhibits the HTLV-1 reactivation from latency via the inhibition of TORC3 function [25]. As outlined in the first chapter of this thesis, Zta is known to interact with TORC2 and this interaction promotes reactivation from latency [90].

First experiments in HEK293 ZKO cells transfected with BCL3 specific siRNA hint at a possible interaction of TORC, BCL3 and Zta. However, while an effect can be observed, it has not been possible to prove a convincing knock-down of BCL3 in these samples as BCL3 expression levels in HEK293 cells are low in comparison to other cell types. In addition, the BCL3 antibody used in earlier western blotting was discontinued and it has not been possible to find a satisfying replacement during the course of this PhD. Future work should clarify the knock-down of BCL3 in the relevant samples.

Unlike Zta, Tax is not a bZIP protein itself, but promotes dimerization of bZIP proteins in the absence of DNA. The elevated concentration of the bZIP homodimer facilitates DNA binding and activation of target sites [173]. TORCs have a conserved N-terminal coiled-coil domain which allows them to interact with the bZIP domain of CREB [174], leading to the induction of cAMP responsive genes. As Zta also contains a bZIP domain [72], it is possible that Zta and CREB interact with TORC in a similar fashion. Furthermore, Zta does not form heterodimers with other bZIP proteins [81], so it is possible that CREB and Zta are competing for the same binding sites. Zta has been shown to decrease transactivation of CREB target sites. This effect is relieved by CBP over-expression, but it has been suggested that Zta does not decrease the amount of CREB-CBP complexes [97].

Knock-down of TORC2 has been shown to reduce the levels of lytic proteins following induction, including Zta [90]. This result was confirmed in this thesis. CBP has been suggested to stabilize a complex containing CREB and HTLV-1 Tax to activate gene transcription. The binding of BCL3 to the CREB/Tax complex inhibits the binding of CBP and represses gene activation. Interestingly, CBP was one of the top five most significant co-associated cellular factor found when intersecting ChIP sequencing MACS peaks, available from ENCODE, with the overlapping Zta ChIP sequencing peaks of HONE1-EBV cells and Akata cells. Knock-down of CBP needs to be validated in further experiments, but treatment of HEK293 ZKO cells with CBP specific siRNA leads to a reduction in Zta protein levels in a similar range to knock-down of TORC1- 3. Further research is needed to examine whether findings relating to CREB, TORC and BCL3 obtained for HTLV-1 Tax can be extended to EBV Zta, and perhaps other lymphotropic viruses like HIV.

ZRE2 contains a CREB binding region and herpes simplex virus type 1 infection has been shown to activate EBV lytic cycle via a CREB dependant mechanism [175]. Perhaps the activation of the B cell receptor, which leads to the activation of CREB's transcriptional activity, results in the induction of the EBV lytic cycle by activating target sites through a similar mechanism later utilised by Zta.

Based on the high overlap of Zta and BCL3 peaks with H3K27me3, the ability of BCL3 to recruit histone deacetylases that inhibit transcription [176] and the fact that Zta has been shown to significantly increases CBP-mediated acetylation of nucleosomal histones [99], a further possible model includes BCL3 involved in the establishment of a repressive chromatin environment during latency, which is overturned through Zta expression and activation at onset of lytic cycle.

Cells expressing the latent EBV protein LMP1 have increased levels of phosphorylated STAT3, resulting in induced BCL3 expression [168]. The expression of LMP1 is thought to result in the p50/BCL3 complex which is the major form for NFκB detected in Nasopharyngeal Carcinoma [169]. In this thesis, results indicate that BCL3 protein levels in EBV-associated cancer cell lines may also increase in response to EBV lytic cycle activation. The constitutive expression of BCL3 has been shown to suppress DNA damage-induced p53 activation and consequent p53-induced apoptosis [159]. BCL3 also inhibits apoptosis by blocking Bim activation [160]. The over-expression of BCL3 has been shown to contribute to the development of some types of Hodgkin and non-Hodgkin lymphoma, as well as nasopharyngeal carcinoma and ovarian cancer [143].

The success of EBV associated cancer therapies is directly linked to the susceptibility for lytic cycle activation. In addition to the anti-apoptotic effect of BCL3, results presented in this thesis raise the question whether increased BCL3 levels are also responsible for the tight latency of EBV associated cancers.

BCL3 has previously been suggested to be a useful diagnostic or prognostic marker for different types of cancers including colorectal cancer [177]. Research presented in the thesis suggests that BCL3 may also be able to fulfil this role for EBV associated cancers. In addition, research at the University of Cardiff has investigated the potential for BCL3 as a new, alternative therapeutic target for HER-2 positive breast cancer [178] and research led by Dr Dafydd Jones at Cardiff University is currently in the process of attempting to develop BCL3 targeted cancer therapies. While more extensive research is required to investigate the model of lytic cycle activation presented in this thesis, BCL3-

targeted cancer therapies could have the potential to improve treatment outcomes for patients with EBV associated cancers.

Bibliography

1. Fields, B.N., D.M. Knipe, and P.M. Howley, *Fields virology*. 5th ed. 2007, Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.
2. Bauman, R.W., E. Machunis-Masuoka, and I.R. Tizard. *Microbiology*. 2004; xxxvii, 780, 88 p. ill. 29 cm. + 1 CD-ROM (4 3/4 in.)].
3. Collier, L.H., P. Kellam, and J.S. Oxford, *Human virology*. 4th ed. 2011, Oxford ; New York: Oxford University Press. xiv, 365 p.
4. Mathers, C.D., T. Boerma, and D. Ma Fat, *Global and regional causes of death.* Br Med Bull, 2009. **92**: p. 7-32.
5. Madigan, M.T., J.M. Martinko, and J. Parker, *Brock biology of microorganisms*. 10th ed. 2003, Upper Saddle River, NJ: Prentice Hall/Pearson Education.
6. Kutok, J.L. and F. Wang, *Spectrum of Epstein-Barr virus-associated diseases.* Annu Rev Pathol, 2006. **1**: p. 375-404.
7. Epstein, M.A., B.G. Achong, and Y.M. Barr, *Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma.* Lancet, 1964. **1**(7335): p. 702-3.
8. Black, F.L., et al., *Prevalence of antibody against viruses in the Tiriyo, an isolated Amazon tribe.* Am J Epidemiol, 1970. **91**(4): p. 430-8.
9. Tischendorf, P., et al., *Development and persistence of immunity to Epstein-Barr virus in man.* J Infect Dis, 1970. **122**(5): p. 401-9.
10. Holmes, D., *The cancer-virus cures.* Nat Med, 2014. **20**(6): p. 571-4.
11. Plummer, M., et al., *Global burden of cancers attributable to infections in 2012: a synthetic analysis.* Lancet Glob Health, 2016. **4**(9): p. e609-16.
12. International Agency for Research on Cancer. and IARC Working Group on the Evaluation of Carcinogenic Risks to Humans., *IARC monographs on the evaluation of carcinogenic risks to humans*. 1988, IARC : Distributed for the International Agency for Research on Cancer by the Secretariat of the World Health Organization: Lyon, France. p. volumes.
13. Pender, M.P. and S.R. Burrows, *Epstein-Barr virus and multiple sclerosis: potential opportunities for immunotherapy.* Clin Transl Immunology, 2014. **3**(10): p. e27.
14. Chan, K.H., et al., *Epstein-Barr virus (EBV) infection in infancy.* J Clin Virol, 2001. **21**(1): p. 57-62.
15. Young, L.S. and A.B. Rickinson, *Epstein-Barr virus: 40 years on.* Nat Rev Cancer, 2004. **4**(10): p. 757-68.
16. Frenette, P., et al., *Factors affecting the age at diagnosis of autism spectrum disorders in Nova Scotia, Canada.* Autism, 2013. **17**(2): p. 184-95.
17. Bollard, C.M., C.M. Rooney, and H.E. Heslop, *T-cell therapy in the treatment of post-transplant lymphoproliferative disease.* Nat Rev Clin Oncol, 2012. **9**(9): p. 510-9.
18. Ramasubramanyan, S., et al., *Genome-wide analyses of Zta binding to the Epstein-Barr virus genome reveals interactions in both early and late lytic cycles and an epigenetic switch leading to an altered binding profile.* J Virol, 2012. **86**(23): p. 12494-502.

19. Sample, J., et al., *Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes.* J Virol, 1990. **64**(9): p. 4084-92.
20. Aitken, C., et al., *Heterogeneity within the Epstein-Barr virus nuclear antigen 2 gene in different strains of Epstein-Barr virus.* J Gen Virol, 1994. **75 (Pt 1)**: p. 95-100.
21. Tao, Q., et al., *Epstein-Barr virus (EBV) and its associated human cancers--genetics, epigenetics, pathobiology and novel therapeutics.* Front Biosci, 2006. **11**: p. 2672-713.
22. Yoon, M., et al., *Mutations in the N termini of herpes simplex virus type 1 and 2 gDs alter functional interactions with the entry/fusion receptors HVEM, nectin-2, and 3-O-sulfated heparan sulfate but not with nectin-1.* J Virol, 2003. **77**(17): p. 9221-31.
23. Silva, A.L., et al., *Mutational analyses of Epstein-Barr virus glycoprotein 42 reveal functional domains not involved in receptor binding but required for membrane fusion.* J Virol, 2004. **78**(11): p. 5946-56.
24. Connolly, S.A., et al., *Fusing structure and function: a structural view of the herpesvirus entry machinery.* Nat Rev Microbiol, 2011. **9**(5): p. 369-81.
25. Shah, K.M. and L.S. Young, *Epstein-Barr virus and carcinogenesis: beyond Burkitt's lymphoma.* Clin Microbiol Infect, 2009. **15**(11): p. 982-8.
26. Kelly, G., A. Bell, and A. Rickinson, *Epstein-Barr virus-associated Burkitt lymphomagenesis selects for downregulation of the nuclear antigen EBNA2.* Nat Med, 2002. **8**(10): p. 1098-104.
27. Faulkner, G.C., et al., *X-Linked agammaglobulinemia patients are not infected with Epstein-Barr virus: implications for the biology of the virus.* J Virol, 1999. **73**(2): p. 1555-64.
28. Wang, D., D. Liebowitz, and E. Kieff, *An EBV membrane protein expressed in immortalized lymphocytes transforms established rodent cells.* Cell, 1985. **43**(3 Pt 2): p. 831-40.
29. Kaye, K.M., K.M. Izumi, and E. Kieff, *Epstein-Barr virus latent membrane protein 1 is essential for B-lymphocyte growth transformation.* Proc Natl Acad Sci U S A, 1993. **90**(19): p. 9150-4.
30. Rowe, M., et al., *Differences in B cell growth phenotype reflect novel patterns of Epstein-Barr virus latent gene expression in Burkitt's lymphoma cells.* EMBO J, 1987. **6**(9): p. 2743-51.
31. Feederle, R., et al., *The Epstein-Barr virus lytic program is controlled by the co-operative functions of two transactivators.* EMBO J, 2000. **19**(12): p. 3080-9.
32. Shannon-Lowe, C. and M. Rowe, *Epstein-Barr virus infection of polarized epithelial cells via the basolateral surface by memory B cell-mediated transfer infection.* PLoS Pathog, 2011. **7**(5): p. e1001338.
33. Borza, C.M. and L.M. Hutt-Fletcher, *Alternate replication in B cells and epithelial cells switches tropism of Epstein-Barr virus.* Nat Med, 2002. **8**(6): p. 594-9.
34. Jiang, R., et al., *Oral dysplasia and squamous cell carcinoma: correlation between increased expression of CD21, Epstein-Barr virus and CK19.* Oral Oncol, 2012. **48**(9): p. 836-41.
35. Jiang, R., et al., *Laser-capture microdissection of oropharyngeal epithelium indicates restriction of Epstein-Barr virus receptor/CD21 mRNA to tonsil epithelial cells.* J Oral Pathol Med, 2008. **37**(10): p. 626-33.
36. Tugizov, S.M., J.W. Berline, and J.M. Palefsky, *Epstein-Barr virus infection of polarized tongue and nasopharyngeal epithelial cells.* Nat Med, 2003. **9**(3): p. 307-14.
37. Xiao, J., et al., *The Epstein-Barr virus BMRF-2 protein facilitates virus attachment to oral epithelial cells.* Virology, 2008. **370**(2): p. 430-42.
38. Chen, J., et al., *Ephrin receptor A2 is a functional entry receptor for Epstein-Barr virus.* Nat Microbiol, 2018. **3**(2): p. 172-180.

39. Shannon-Lowe, C., et al., *Features distinguishing Epstein-Barr virus infections of epithelial cells and B cells: viral genome expression, genome maintenance, and genome amplification.* J Virol, 2009. **83**(15): p. 7749-60.
40. Pfeffer, S., et al., *Identification of virus-encoded microRNAs.* Science, 2004. **304**(5671): p. 734-6.
41. Grundhoff, A., C.S. Sullivan, and D. Ganem, *A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses.* RNA, 2006. **12**(5): p. 733-50.
42. Cai, X., et al., *Epstein-Barr virus microRNAs are evolutionarily conserved and differentially expressed.* PLoS Pathog, 2006. **2**(3): p. e23.
43. Cochet, C., et al., *Expression of the Epstein-Barr virus immediate early gene, BZLF1, in nasopharyngeal carcinoma tumor cells.* Virology, 1993. **197**(1): p. 358-65.
44. Raab-Traub, N. and K. Flynn, *The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation.* Cell, 1986. **47**(6): p. 883-9.
45. Gulley, M.L., et al., *Epstein-Barr virus is detected in undifferentiated nasopharyngeal carcinoma but not in lymphoepithelioma-like carcinoma of the urinary bladder.* Hum Pathol, 1995. **26**(11): p. 1207-14.
46. Low, W.K., et al., *Diagnostic value of Epstein-Barr viral serology in nasopharyngeal carcinoma.* Otolaryngol Head Neck Surg, 2000. **123**(4): p. 505-7.
47. Henle, G. and W. Henle, *Epstein-Barr virus-specific IgA serum antibodies as an outstanding feature of nasopharyngeal carcinoma.* Int J Cancer, 1976. **17**(1): p. 1-7.
48. Chan, A.T., et al., *Plasma Epstein-Barr virus DNA and residual disease after radiotherapy for undifferentiated nasopharyngeal carcinoma.* J Natl Cancer Inst, 2002. **94**(21): p. 1614-9.
49. Temple, R.M., et al., *Efficient replication of Epstein-Barr virus in stratified epithelium in vitro.* Proc Natl Acad Sci U S A, 2014. **111**(46): p. 16544-9.
50. Sbih-Lammali, F., et al., *Transcriptional expression of Epstein-Barr virus genes and proto-oncogenes in north African nasopharyngeal carcinoma.* J Med Virol, 1996. **49**(1): p. 7-14.
51. Young, L.S., L.F. Yap, and P.G. Murray, *Epstein-Barr virus: more than 50 years old and still providing surprises.* Nat Rev Cancer, 2016. **16**(12): p. 789-802.
52. Hutajulu, S.H., et al., *Epigenetic markers for early detection of nasopharyngeal carcinoma in a high risk population.* Mol Cancer, 2011. **10**: p. 48.
53. Knox, P.G., et al., *In vitro production of stable Epstein-Barr virus-positive epithelial cell clones which resemble the virus:cell interaction observed in nasopharyngeal carcinoma.* Virology, 1996. **215**(1): p. 40-50.
54. Caldwell, R.G., et al., *Epstein-Barr virus LMP2A drives B cell development and survival in the absence of normal B cell receptor signals.* Immunity, 1998. **9**(3): p. 405-11.
55. Hong, G.K., et al., *Epstein-Barr virus lytic infection is required for efficient production of the angiogenesis factor vascular endothelial growth factor in lymphoblastoid cell lines.* J Virol, 2005. **79**(22): p. 13984-92.
56. Klein, G., et al., *Direct evidence for the presence of Epstein-Barr virus DNA and nuclear antigen in malignant epithelial cells from patients with poorly differentiated carcinoma of the nasopharynx.* Proc Natl Acad Sci U S A, 1974. **71**(12): p. 4737-41.
57. Niedobitek, G., et al., *Epstein-Barr virus and carcinomas: undifferentiated carcinomas but not squamous cell carcinomas of the nasopharynx are regularly associated with the virus.* J Pathol, 1991. **165**(1): p. 17-24.
58. Pathmanathan, R., et al., *Undifferentiated, nonkeratinizing, and squamous cell carcinoma of the nasopharynx. Variants of Epstein-Barr virus-infected neoplasia.* Am J Pathol, 1995. **146**(6): p. 1355-67.

59. Pathmanathan, R., et al., *Clonal proliferations of cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal carcinoma.* N Engl J Med, 1995. **333**(11): p. 693-8.
60. Niedobitek, G., et al., *Expression of Epstein-Barr virus genes and of lymphocyte activation molecules in undifferentiated nasopharyngeal carcinomas.* Am J Pathol, 1992. **140**(4): p. 879-87.
61. Heussinger, N., et al., *Expression of the Epstein-Barr virus (EBV)-encoded latent membrane protein 2A (LMP2A) in EBV-associated nasopharyngeal carcinoma.* J Pathol, 2004. **203**(2): p. 696-9.
62. Young, L.S., et al., *Epstein-Barr virus gene expression in nasopharyngeal carcinoma.* J Gen Virol, 1988. **69 (Pt 5)**: p. 1051-65.
63. Sokal, E.M., et al., *Recombinant gp350 vaccine for infectious mononucleosis: a phase 2, randomized, double-blind, placebo-controlled trial to evaluate the safety, immunogenicity, and efficacy of an Epstein-Barr virus vaccine in healthy young adults.* J Infect Dis, 2007. **196**(12): p. 1749-53.
64. Hjalgrim, H., et al., *Characteristics of Hodgkin's lymphoma after infectious mononucleosis.* N Engl J Med, 2003. **349**(14): p. 1324-32.
65. Handel, A.E. and S.V. Ramagopalan, *Multiple sclerosis and risk of cancer: a meta-analysis.* J Neurol Neurosurg Psychiatry, 2010. **81**(12): p. 1413-4.
66. Coghill, A.E., et al., *High Levels of Antibody that Neutralize B-cell Infection of Epstein-Barr Virus and that Bind EBV gp350 Are Associated with a Lower Risk of Nasopharyngeal Carcinoma.* Clin Cancer Res, 2016. **22**(14): p. 3451-7.
67. deBruyn, J.C.C., et al., *Serologic Status of Routine Childhood Vaccines, Cytomegalovirus, and Epstein-Barr Virus in Children With Inflammatory Bowel Disease.* Inflamm Bowel Dis, 2018.
68. Cohen, J.I., *Vaccine Development for Epstein-Barr Virus.* Adv Exp Med Biol, 2018. **1045**: p. 477-493.
69. Pudney, V.A., et al., *CD8+ immunodominance among Epstein-Barr virus lytic cycle antigens directly reflects the efficiency of antigen presentation in lytically infected cells.* J Exp Med, 2005. **201**(3): p. 349-60.
70. Brooks, J.M., et al., *Early T Cell Recognition of B Cells following Epstein-Barr Virus Infection: Identifying Potential Targets for Prophylactic Vaccination.* PLoS Pathog, 2016. **12**(4): p. e1005549.
71. Marschall, M., et al., *Identification of proteins encoded by Epstein-Barr virus trans-activator genes.* J Virol, 1989. **63**(2): p. 938-42.
72. Petosa, C., et al., *Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein.* Mol Cell, 2006. **21**(4): p. 565-72.
73. El-Guindy, A., et al., *Essential role of Rta in lytic DNA replication of Epstein-Barr virus.* J Virol, 2013. **87**(1): p. 208-23.
74. Countryman, J. and G. Miller, *Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a small cloned subfragment of heterogeneous viral DNA.* Proc Natl Acad Sci U S A, 1985. **82**(12): p. 4085-9.
75. Kalla, M., et al., *AP-1 homolog BZLF1 of Epstein-Barr virus has two essential functions dependent on the epigenetic state of the viral genome.* Proc Natl Acad Sci U S A, 2010. **107**(2): p. 850-5.
76. Ramasubramanyan, S., et al., *Dynamic chromatin environment of key lytic cycle regulatory regions of the Epstein-Barr virus genome.* J Virol, 2012. **86**(3): p. 1809-19.
77. Fernandez, A.F., et al., *The dynamic DNA methylomes of double-stranded DNA viruses associated with human cancer.* Genome Res, 2009. **19**(3): p. 438-51.
78. Bergbauer, M., et al., *CpG-methylation regulates a class of Epstein-Barr virus promoters.* PLoS Pathog, 2010. **6**(9): p. e1001114.

79. Flower, K., et al., *Epigenetic control of viral life-cycle by a DNA-methylation dependent transcription factor.* PLoS One, 2011. **6**(10): p. e25922.
80. Sinclair, A.J., *bZIP proteins of human gammaherpesviruses.* J Gen Virol, 2003. **84**(Pt 8): p. 1941-9.
81. Chang, Y.N., et al., *The Epstein-Barr virus Zta transactivator: a member of the bZIP family with unique DNA-binding specificity and a dimerization domain that lacks the characteristic heptad leucine zipper motif.* J Virol, 1990. **64**(7): p. 3358-69.
82. Gutsch, D.E., et al., *The bZIP transactivator of Epstein-Barr virus, BZLF1, functionally and physically interacts with the p65 subunit of NF-kappa B.* Mol Cell Biol, 1994. **14**(3): p. 1939-48.
83. Sista, N.D., et al., *Retinoic acid is a negative regulator of the Epstein-Barr virus protein (BZLF1) that mediates disruption of latent infection.* Proc Natl Acad Sci U S A, 1993. **90**(9): p. 3894-8.
84. Zhang, Q., D. Gutsch, and S. Kenney, *Functional and physical interaction between p53 and BZLF1: implications for Epstein-Barr virus latency.* Mol Cell Biol, 1994. **14**(3): p. 1929-38.
85. Lin, J.C., M.C. Smith, and J.S. Pagano, *Induction of replication of Epstein-Barr virus DNA by 12-O-tetradecanoyl-phorbol-13-acetate. II. Inhibition by retinoic acid and 9-(2-hydroxyethoxymethyl) guanine.* Virology, 1981. **111**(1): p. 294-8.
86. Yamamoto, N., K. Bister, and H. zur Hausen, *Retinoic acid inhibition of Epstein-Barr virus induction.* Nature, 1979. **278**(5704): p. 553-4.
87. Lieberman, P.M. and A.J. Berk, *The Zta trans-activator protein stabilizes TFIID association with promoter DNA by direct protein-protein interaction.* Genes Dev, 1991. **5**(12B): p. 2441-54.
88. Gao, Z., et al., *The Epstein-Barr virus lytic transactivator Zta interacts with the helicase-primase replication proteins.* J Virol, 1998. **72**(11): p. 8559-67.
89. Lieberman, P.M., et al., *The zta transactivator involved in induction of lytic cycle gene expression in Epstein-Barr virus-infected lymphocytes binds to both AP-1 and ZRE sites in target promoter and enhancer regions.* J Virol, 1990. **64**(3): p. 1143-55.
90. Murata, T., et al., *TORC2, a coactivator of cAMP-response element-binding protein, promotes Epstein-Barr virus reactivation from latency through interaction with viral BZLF1 protein.* J Biol Chem, 2009. **284**(12): p. 8033-41.
91. Jamaluddin, M., et al., *Respiratory syncytial virus-inducible BCL-3 expression antagonizes the STAT/IRF and NF-kappaB signaling pathways by inducing histone deacetylase 1 recruitment to the interleukin-8 promoter.* J Virol, 2005. **79**(24): p. 15302-13.
92. Liang, C.L., et al., *Epstein-Barr virus BZLF1 gene is activated by transforming growth factor-beta through cooperativity of Smads and c-Jun/c-Fos proteins.* J Biol Chem, 2002. **277**(26): p. 23345-57.
93. Wang, Y.G., et al., *Targeted disruption of the CREB coactivator Crtc2 increases insulin sensitivity.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(7): p. 3087-3092.
94. Altarejos, J.Y. and M. Montminy, *CREB and the CRTC co-activators: sensors for hormonal and metabolic signals.* Nat Rev Mol Cell Biol, 2011. **12**(3): p. 141-51.
95. Ravnskjaer, K., et al., *Cooperative interactions between CBP and TORC2 confer selectivity to CREB target gene expression.* Embo Journal, 2007. **26**(12): p. 2880-2889.
96. Swenson, J.J., E. Holley-Guthrie, and S.C. Kenney, *Epstein-Barr virus immediate-early protein BRLF1 interacts with CBP, promoting enhanced BRLF1 transactivation.* J Virol, 2001. **75**(13): p. 6228-34.
97. Adamson, A.L. and S. Kenney, *The Epstein-Barr virus BZLF1 protein interacts physically and functionally with the histone acetylase CREB-binding protein.* J Virol, 1999. **73**(8): p. 6551-8.

98. Zerby, D., et al., *The amino-terminal C/H1 domain of CREB binding protein mediates zta transcriptional activation of latent Epstein-Barr virus.* Mol Cell Biol, 1999. **19**(3): p. 1617-26.
99. Chen, C.J., et al., *Stimulation of CREB binding protein nucleosomal histone acetyltransferase activity by a class of transcriptional activators.* Mol Cell Biol, 2001. **21**(2): p. 476-87.
100. Kamei, Y., et al., *A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors.* Cell, 1996. **85**(3): p. 403-14.
101. Zhong, H., R.E. Voll, and S. Ghosh, *Phosphorylation of NF-kappa B p65 by PKA stimulates transcriptional activity by promoting a novel bivalent interaction with the coactivator CBP/p300.* Mol Cell, 1998. **1**(5): p. 661-71.
102. Strong, M.J., et al., *Comprehensive high-throughput RNA sequencing analysis reveals contamination of multiple nasopharyngeal carcinoma cell lines with HeLa cell genomes.* J Virol, 2014. **88**(18): p. 10696-704.
103. Lui, V.W., et al., *STAT3 activation contributes directly to Epstein-Barr virus-mediated invasiveness of nasopharyngeal cancer cells in vitro.* Int J Cancer, 2009. **125**(8): p. 1884-93.
104. Tsai, M.H., et al., *Spontaneous lytic replication and epitheliotropism define an Epstein-Barr virus strain found in carcinomas.* Cell Rep, 2013. **5**(2): p. 458-70.
105. Ramasubramanyan, S., et al., *Epstein-Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression.* Nucleic Acids Res, 2015.
106. DuBridge, R.B., et al., *Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system.* Mol Cell Biol, 1987. **7**(1): p. 379-87.
107. Yao, K.T., et al., *Establishment and characterization of two epithelial tumor cell lines (HNE-1 and HONE-1) latently infected with Epstein-Barr virus and derived from nasopharyngeal carcinomas.* Int J Cancer, 1990. **45**(1): p. 83-9.
108. Sinclair, A.J., et al., *EBNA-2 and EBNA-LP cooperate to cause G0 to G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus.* EMBO J, 1994. **13**(14): p. 3321-8.
109. Young, L.S., et al., *Differentiation-associated expression of the Epstein-Barr virus BZLF1 transactivator protein in oral hairy leukoplakia.* J Virol, 1991. **65**(6): p. 2868-74.
110. Hollyoake, M., et al., *The normal cell cycle activation program is exploited during the infection of quiescent B lymphocytes by Epstein-Barr virus.* Cancer Res, 1995. **55**(21): p. 4784-7.
111. Ramasubramanyan, S., et al., *Epstein-Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression.* Nucleic Acids Res, 2015. **43**(7): p. 3563-77.
112. Gallagher, A., et al., *Detection of Epstein-Barr virus (EBV) genomes in the serum of patients with EBV-associated Hodgkin's disease.* Int J Cancer, 1999. **84**(4): p. 442-8.
113. Palermo, R.D., H.M. Webb, and M.J. West, *RNA polymerase II stalling promotes nucleosome occlusion and pTEFb recruitment to drive immortalization by Epstein-Barr virus.* PLoS Pathog, 2011. **7**(10): p. e1002334.
114. Brenne, A.T., et al., *High expression of BCL3 in human myeloma cells is associated with increased proliferation and inferior prognosis.* Eur J Haematol, 2009. **82**(5): p. 354-63.
115. Bailey, S.G., et al., *Functional interaction between Epstein-Barr virus replication protein Zta and host DNA damage response protein 53BP1.* J Virol, 2009. **83**(21): p. 11116-22.
116. Ramirez, F., et al., *deepTools: a flexible platform for exploring deep-sequencing data.* Nucleic Acids Res, 2014. **42**(Web Server issue): p. W187-91.
117. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Res, 2002. **12**(6): p. 996-1006.

118. Machanick, P. and T.L. Bailey, *MEME-ChIP: motif analysis of large DNA datasets.* Bioinformatics, 2011. **27**(12): p. 1696-7.
119. Chen, E.Y., et al., *Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool.* BMC Bioinformatics, 2013. **14**: p. 128.
120. McLean, C.Y., et al., *GREAT improves functional interpretation of cis-regulatory regions.* Nat Biotechnol, 2010. **28**(5): p. 495-501.
121. Coletta, A., et al., *InSilico DB genomic datasets hub: an efficient starting point for analyzing genome-wide studies in GenePattern, Integrative Genomics Viewer, and R/Bioconductor.* Genome Biol, 2012. **13**(11): p. R104.
122. Ye, F., et al., *Genetic profiling reveals an alarming rate of cross-contamination among human cell lines used in China.* FASEB J, 2015. **29**(10): p. 4268-72.
123. Chan, S.Y., et al., *Authentication of nasopharyngeal carcinoma tumor lines.* Int J Cancer, 2008. **122**(9): p. 2169-71.
124. Dokmanovic, M., C. Clarke, and P.A. Marks, *Histone deacetylase inhibitors: overview and perspectives.* Mol Cancer Res, 2007. **5**(10): p. 981-9.
125. Bubna, A.K., *Vorinostat-An Overview.* Indian J Dermatol, 2015. **60**(4): p. 419.
126. Soto Chervin, C. and B. Brockstein, *Current clinical immunotherapeutic approaches for head and neck cancer.* F1000Res, 2016. **5**.
127. Miller, G., *The switch between latency and replication of Epstein-Barr virus.* J Infect Dis, 1990. **161**(5): p. 833-44.
128. Lu, C.C., et al., *Genome-wide transcription program and expression of the Rta responsive gene of Epstein-Barr virus.* Virology, 2006. **345**(2): p. 358-72.
129. Vogler, M., *BCL2A1: the underdog in the BCL2 family.* Cell Death Differ, 2012. **19**(1): p. 67-74.
130. Hutt-Fletcher, L.M., *Epstein-Barr virus replicating in epithelial cells.* Proc Natl Acad Sci U S A, 2014. **111**(46): p. 16242-3.
131. Kuleshov, M.V., et al., *Enrichr: a comprehensive gene set enrichment analysis web server 2016 update.* Nucleic Acids Res, 2016. **44**(W1): p. W90-7.
132. Machanick, P. and T.L. Bailey, *MEME-ChIP: motif analysis of large DNA datasets.* Bioinformatics, 2011. **27**(12): p. 1696-1697.
133. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.
134. Diaz, A., et al., *Normalization, bias correction, and peak calling for ChIP-seq.* Statistical Applications in Genetics and Molecular Biology, 2012. **11**(3).
135. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome.* Cell, 2007. **129**(4): p. 823-37.
136. Thomas, P.D., et al., *PANTHER: a library of protein families and subfamilies indexed by function.* Genome Res, 2003. **13**(9): p. 2129-41.
137. Gene Ontology, C., *Gene Ontology Consortium: going forward.* Nucleic Acids Res, 2015. **43**(Database issue): p. D1049-56.
138. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions.* Nat Rev Genet, 2014. **15**(4): p. 272-86.
139. Bailey, T., et al., *Practical guidelines for the comprehensive analysis of ChIP-seq data.* PLoS Comput Biol, 2013. **9**(11): p. e1003326.
140. Simon, J.A. and R.E. Kingston, *Mechanisms of polycomb gene silencing: knowns and unknowns.* Nat Rev Mol Cell Biol, 2009. **10**(10): p. 697-708.
141. Brinkman, A.B., et al., *Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk.* Genome Res, 2012. **22**(6): p. 1128-38.
142. Statham, A.L., et al., *Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA.* Genome Res, 2012. **22**(6): p. 1120-7.

143. Maldonado, V. and J. Melendez-Zajgla, *Role of Bcl-3 in solid tumors.* Mol Cancer, 2011. **10**: p. 152.
144. Schwab, M., *Encyclopedia of cancer*. 3rd ed. Springer reference. 2011, Heidelberg ; New York: Springer.
145. Zerbino, D.R., et al., *Ensembl 2018.* Nucleic Acids Res, 2018. **46**(D1): p. D754-D761.
146. Elliott, W.H. and D.C. Elliott, *Biochemistry and molecular biology*. 3rd ed. 2005, Oxford ; New York: Oxford University Press. xxxiii, 582 p.
147. Oeckinghaus, A. and S. Ghosh, *The NF-kappaB family of transcription factors and its regulation.* Cold Spring Harb Perspect Biol, 2009. **1**(4): p. a000034.
148. Cristofanon, S., et al., *Oxidative, multistep activation of the noncanonical NF-kappaB pathway via disulfide Bcl-3/p50 complex.* FASEB J, 2009. **23**(1): p. 45-57.
149. Michel, F., et al., *Crystal structure of the ankyrin repeat domain of Bcl-3: a unique member of the IkappaB protein family.* EMBO J, 2001. **20**(22): p. 6180-90.
150. Wessells, J., et al., *BCL-3 and NF-kappaB p50 attenuate lipopolysaccharide-induced inflammatory responses in macrophages.* J Biol Chem, 2004. **279**(48): p. 49995-50003.
151. Carmody, R.J., et al., *Negative regulation of toll-like receptor signaling by NF-kappaB p50 ubiquitination blockade.* Science, 2007. **317**(5838): p. 675-8.
152. Gaudet, P., et al., *The neXtProt knowledgebase on human proteins: current status.* Nucleic Acids Res, 2015. **43**(Database issue): p. D764-70.
153. Bredemeyer, A.L., et al., *DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes.* Nature, 2008. **456**(7223): p. 819-23.
154. Na, S.Y., et al., *Bcl3, an IkappaB protein, stimulates activating protein-1 transactivation and cellular proliferation.* J Biol Chem, 1999. **274**(40): p. 28491-6.
155. Dechend, R., et al., *The Bcl-3 oncoprotein acts as a bridging factor between NF-kappaB/Rel and nuclear co-regulators.* Oncogene, 1999. **18**(22): p. 3316-23.
156. Hishiki, T., et al., *BCL3 acts as a negative regulator of transcription from the human T-cell leukemia virus type 1 long terminal repeat through interactions with TORC3.* J Biol Chem, 2007. **282**(39): p. 28335-43.
157. Goncalves, D.U., et al., *Epidemiology, treatment, and prevention of human T-cell leukemia virus type 1-associated diseases.* Clin Microbiol Rev, 2010. **23**(3): p. 577-89.
158. Kim, Y.M., N. Sharma, and J.K. Nyborg, *The proto-oncogene Bcl3, induced by Tax, represses Tax-mediated transcription via p300 displacement from the human T-cell leukemia virus type 1 promoter.* J Virol, 2008. **82**(23): p. 11939-47.
159. Kashatus, D., P. Cogswell, and A.S. Baldwin, *Expression of the Bcl-3 proto-oncogene suppresses p53 activation.* Genes Dev, 2006. **20**(2): p. 225-35.
160. Bauer, A., et al., *The NF-kappaB regulator Bcl-3 and the BH3-only proteins Bim and Puma control the death of activated T cells.* Proc Natl Acad Sci U S A, 2006. **103**(29): p. 10979-84.
161. Fornari, T.A., et al., *Development of type 1 diabetes mellitus in nonobese diabetic mice follows changes in thymocyte and peripheral T lymphocyte transcriptional activity.* Clin Dev Immunol, 2011. **2011**: p. 158735.
162. Massoumi, R., *The central role of Bcl-3 in atopic dermatitis.* J Invest Dermatol, 2009. **129**(9): p. 2088-90.
163. Reissig, S., et al., *Elevated levels of Bcl-3 inhibits Treg development and function resulting in spontaneous colitis.* Nat Commun, 2017. **8**: p. 15069.
164. Guan, Y., et al., *MiR-125b targets BCL3 and suppresses ovarian cancer proliferation.* Int J Cancer, 2011. **128**(10): p. 2274-83.
165. Ong, S.T., et al., *Lymphadenopathy, splenomegaly, and altered immunoglobulin production in BCL3 transgenic mice.* Oncogene, 1998. **16**(18): p. 2333-43.
166. Zhang, X., et al., *The tumor promoter and NF-kappaB modulator Bcl-3 regulates splenic B cell development.* J Immunol, 2013. **191**(12): p. 5984-92.

167. Poljak, L., et al., *Distinct activities of p52/NF-kappa B required for proper secondary lymphoid organ microarchitecture: functions enhanced by Bcl-3.* J Immunol, 1999. **163**(12): p. 6581-8.
168. Thornburg, N.J. and N. Raab-Traub, *Induction of epidermal growth factor receptor expression by Epstein-Barr virus latent membrane protein 1 C-terminal-activating region 1 is mediated by NF-kappaB p50 homodimer/Bcl-3 complexes.* J Virol, 2007. **81**(23): p. 12954-61.
169. Chung, G.T., et al., *Constitutive activation of distinct NF-kappaB signals in EBV-associated nasopharyngeal carcinoma.* J Pathol, 2013. **231**(3): p. 311-22.
170. Takada, K. and Y. Ono, *Synchronous and sequential activation of latently infected Epstein-Barr virus genomes.* J Virol, 1989. **63**(1): p. 445-9.
171. Arvey, A., I. Tempera, and P.M. Lieberman, *Interpreting the Epstein-Barr Virus (EBV) epigenome using high-throughput data.* Viruses, 2013. **5**(4): p. 1042-54.
172. Godfrey, A., S. Ramasubramanyan, and A.J. Sinclair, *The Use of Chromatin Precipitation Coupled to DNA Sequencing (ChIP-Seq) for the Analysis of Zta Binding to the Human and EBV Genome.* Methods Mol Biol, 2017. **1532**: p. 191-206.
173. Wagner, S. and M.R. Green, *HTLV-I Tax protein stimulation of DNA binding of bZIP proteins by enhancing dimerization.* Science, 1993. **262**(5132): p. 395-9.
174. Lever, A.M.L., K.-T. Jeang, and B. Berkhout, *Recent advances in human retroviruses : principles of replication and pathogenesis : advances in retroviral research*. 2010, Singapore ; Hackensack, NJ: World Scientific. vii, 493 p.
175. Wu, H., et al., *Herpes simplex virus type 1 infection activates the Epstein-Barr virus replicative cycle via a CREB-dependent mechanism.* Cell Microbiol, 2012. **14**(4): p. 546-59.
176. Jamaluddin, M., et al., *Respiratory syncytial virus-inducible BCL-3 expression antagonizes the STAT/IRF and NF-kappa B signaling pathways by inducing histone deacetylase 1 recruitment to the interleukin-8 promoter.* Journal of Virology, 2005. **79**(24): p. 15302-15313.
177. Saamarthy, K., et al., *Early diagnostic value of Bcl-3 localization in colorectal cancer.* BMC Cancer, 2015. **15**: p. 341.
178. Wakefield, A., et al., *Bcl3 selectively promotes metastasis of ERBB2-driven mammary tumors.* Cancer Res, 2013. **73**(2): p. 745-55.

Appendix A

Supplementary Figure 1: List of Genes with Zta MACS peaks in HONE1-EBV cells within 2kb of the transcription start site. List was obtained using GREAT version 2.0.2, Species assembly: hg19 and the association rule: Two nearest genes: 2000 bp max extension, curated regulatory domains included.

ABCA12	ASB13	C12orf68	CELSR1
ABCC12	ASCL1	C13orf15	CEP72
ACMSD	ASPH	C15orf44	CGB
ACOT9	ASPHD2	C19orf68	CGB2
ACTA1	ASXL3	C1orf130	CGREF1
ACY3	ATP1A3	C2CD2	CHRNA5
ADAMTS10	ATP8A2	C5orf43	CHST7
ADRM1	ATP9A	CA5A	CHST8
AFAP1L1	ATPAF2	CAB39	CHSY3
AGTPBP1	AUTS2	CACNG3	CLDN12
ALDH3A1	BANP	CACNG4	CNOT1
ALS2	BBC3	CALN1	CNOT10
ALS2CR12	BCAN	CALY	CNTFR
AMOTL1	BCAT1	CAMSAP3	CNTN1
AMT	BCCIP	CBLC	COL27A1
AMZ1	BCL11B	CCAR1	COL5A1
ANKRD54	BCL2A1	CCDC70	COL5A3
ANO3	BCL7C	CD27	COPG
ANXA2	BDKRB1	CD300A	COTL1
AP1S2	BLVRB	CD300C	CPN2
ARHGDIB	BOLA1	CD3E	CPNE6
ARL4C	BRD1	CD97	CPXM1
ARMC7	BTAF1	CDKN1C	CREG2
ARNT	BZRAP1	CDX2	CRMP1
ARRB1	C11orf20	CELF4	CRTC3

CSGALNACT1	DOC2B	FAM8A1	GLIS1
CSH1	DPF1	FBXL14	GNA12
CTF1	DPH3P1	FBXO2	GNG4
CTSW	DSG3	FBXO22	GP1BB
CTSZ	EBNA1BP2	FEM1C	GPATCH4
CYC1	EFNA3	FES	GPC2
CYGB	EHD2	FEZ1	GPR89A
CYP26B1	EID3	FGF13	GPR89B
CYP26C1	EIF2B3	FLT1	GRIN1
CYP4F12	ELFN2	FN1	GSG1
CYP4F8	ELMO3	FNDC4	GTF2IRD1
DACH1	ELN	FOXD3	GYG2
DACH2	EMP1	FOXI1	HAND1
DCHS1	EPHA4	FOXK2	HAS1
DDX27	EPHB3	FSD1	HBB
DEGS2	EPHX3	FZD9	HBD
DENND1C	ERBB4	G3BP1	HBE1
DHRS7C	ERMN	GABRA4	HBG1
DLEC1	ETV3	GBX2	HBZ
DLK1	EXOC6B	GCDH	HEATR4
DLL3	FAM109A	GCFC1	HECW1
DNAJB1	FAM110C	GDF2	HIST1H3J
DNAJC22	FAM184A	GEM	HIST3H3
DNHD1	FAM20C	GGT5	HMGCS1
DOC2A	FAM69A	GIPR	HMGCS2

HMX1	KIFC3	MDC1	NOTCH3
HMX2	KLC3	MDGA2	NOVA1
HNF4A	KLK5	MDM4	NOVA2
HOXA13	KRT18	MEAF6	NPIPL1
HOXA2	KRTAP4-5	MGRN1	NPIPL3
HOXB7	LAMA3	MMP1	NR2C1
HOXB8	LAMB1	MREG	NRBF2
HSPA12B	LCP1	MRPL38	NTN1
IDUA	LGALS9	MRPL41	NTNG1
IGBP1	LIPH	MRPL55	NTRK2
IGDCC3	LMTK3	MRPS17	NUCKS1
IL11RA	LOC100132247	MS4A10	NUP93
IQSEC1	LOC100132247	MYL6	ODZ4
IRGM	LOC613037	MYO1B	OLFM2
ISLR2	LRFN5	MYO5A	ONECUT1
ITFG3	LRIG1	NCKAP1L	OPN4
ITPKB	LRP10	NDUFA6	OR2T2
IYD	LRP2	NDUFC2-KCTD14	OR2T35
KCNC3	LRRC33	NEIL3	P4HA3
KCNG3	LYSMD2	NFE2L3	PCIF1
KCNIP2	MAN2A2	NGLY1	PCNT
KCNN4	MAN2B1	NKX6-2	PDE10A
KCP	MAN2C1	NMNAT2	PDE1B
KIAA0664	MAST1	NMUR1	PDE4A
KIF2A	MATK	NOMO3	PDZK1

PGK1	PSPH	SCIMP	SNRPC
PGM3	PTDSS2	SCRT1	SOCS1
PHF21B	PTGR1	SCT	SORCS3
PLA2G10	R3HDML	SDSL	SORL1
PLA2G16	RAB33A	SEC14L5	SOX3
PLA2G4C	RAB3IP	SEMA3B	SPATA24
PLA2G4F	RANBP3L	SETDB1	SPDEF
PLCD1	RASGEF1A	SETDB2	SPTBN4
PLK5	RASGRF2	SEZ6L	SSH1
PNMA3	RBM27	SFRP1	STAG3
POLA2	RELT	SFRP5	STAT6
POTED	RFFL	SHANK3	STRAP
POU4F1	RFXANK	SHH	SYCP2
POU4F2	RFXAP	SHISA6	SYT12
POU4F3	RIN2	SHROOM1	SYTL3
PPFIA1	RMI2	SLC18A3	TAB1
PPM1K	RNASET2	SLC22A18	TAB2
PRCD	RPF1	SLC2A2	TAS1R3
PRLH	RPL23	SLC45A2	TBKBP1
PRM1	RPLP0	SLC6A3	TCF15
PRM2	RPUSD3	SLC9A11	TECR
PROX1	RRAS2	SLCO2A1	TGFBR3
PRPF3	RRP1	SLITRK5	TH1L
PRUNE	RTN4RL2	SLPI	THBS2
PSMC3	S100A16	SMAGP	THRA

TIMMDC1	TRAPPC6A	UCN2	XKR8
TMEM117	TRIB2	UHRF1BP1	ZCCHC14
TMEM130	TRIM54	UNK	ZFP36
TMEM211	TRIM7	USH1G	ZFYVE26
TMEM26	TRPM7	USP47	ZNF148
TMEM74B	TSSK6	USP54	ZNF205
TMEM85	TSTA3	VSX2	ZNF341
TNFSF14	TTBK2	VWA1	ZNF527
TOX2	TTC9B	WDR34	ZNF554
TPM4	UBE2I	WDR45L	ZNF566
TPT1	UBE2J2	WFIKKN2	ZNF583
TRAF3	UBL3	WNT8A	
TRAPPC5	UBOX5	WSCD1	

Supplementary Figure 2: List of Genes with Zta MACS peaks, common to epithelial and B cells, within 2kb of the transcription start site. List was obtained using GREAT version 2.0.2, Species assembly: hg19 and the association rule: Two nearest genes: 2000 bp max extension, curated regulatory domains included.

ACY3	ARRB1	BCL2A1	C1orf130
ADAMTS10	ATP1A3	BCL7C	CA5A
ALDH3A1	ATPAF2	BLVRB	CALN1
ANKRD54	BANP	BRD1	CALY
ANXA2	BBC3	BZRAP1	CAMSAP3

CD300A	DENND1C	GPATCH4	MGRN1
CD300C	DLK1	GRIN1	MLYCD
CELF4	DOC2A	HAS1	MRPL41
CELSR1	DOC2B	HNF4A	NOVA2
CGB	DPF1	HOXA13	NUMBL
CGB2	DPH3P1	HSPA12B	OLFM2
CGREF1	EFNA3	IDUA	ONECUT1
CNOT10	ELMO3	IL11RA	OPRL1
COL27A1	EPHA4	IRGM	PCIF1
COL5A3	EPHB3	IYD	PCNT
CPNE6	FAIM3	KCNIP2	PDE1B
CPXM1	FAM109A	KIFC3	PDE4A
CRMP1	FAM20C	KLC3	PGK1
CTF1	FAM70B	LAMB1	PHF21B
CYGB	FBXL14	LCP1	PLA2G4C
CYP4F11	FES	LMTK3	PLA2G4F
CYP4F12	GCDH	LRP2	PLCD1
CYP4F8	GDF5OS	MAN2B1	PPIL2
DCHS1	GIPR	MAST1	PRCD
DDX27	GP1BB	MEAF6	PRLH

PTDSS2	SLC45A2	UBOX5
PTK6	SLC6A3	VWA1
R3HDML	SORL1	WDR45L
RAN	SPTBN4	WFIKKN2
RASGRF2	SSH1	WSCD1
RFXANK	TAB1	ZNF205
RPL23	TAS1R3	ZNF554
RRAS2	THRA	ZNF566
RRP1	TNFSF14	ZNF583
RTN4RL2	TOX2	
SCIMP	TPM4	
SCT	TPT1	
SEC14L5	TRAF3	
SEMA3B	TRAPPC5	
SEZ6L	TRIM8	
SHANK3	TSSK6	
SHH	TSTA3	
SLC22A18	TTC9B	
SLC22A20	UBE2I	
SLC27A1	UBE2J2	

Appendix B

Chapter 14

The Use of Chromatin Precipitation Coupled to DNA Sequencing (ChIP-Seq) for the Analysis of Zta Binding to the Human and EBV Genome

Anja Godfrey, Sharada Ramasubramanyan, and Alison J. Sinclair

Abstract

Determining which components of the transcription machinery associate with the viral and cellular genome, and how this changes at specific stages of the viral life cycle is paramount to understanding how the distinct transcriptional programs associated with primary infection, latency, and disease are established and how they are reprogrammed during initiation and execution of the viral lytic replication cycle. Chromatin precipitations linked to next generation DNA sequencing (ChIP-Seq) allow for the interactions of proteins with DNA to be mapped across both viral and cellular genomes. This can be applied to viral and cellular transcription factors, coactivators and corepressors, modified histones, and modulators of chromatin.

Key words Chromatin, Immunoprecipitation, DNA, ChIP, DNA sequencing, Transcription factor

1 Introduction

Every cell in an organism has an identical genome, which contains all of the information required for the lifetime of that organism, but many genes are required only in specialized cells or in response to particular conditions. Highly controlled mechanisms that regulate the expression of such genes are paramount to the specialization of cells, thus allowing for the complexity and survival of the organism. This is achieved through a myriad of controls that ensure that an environment that is either favorable or unfavorable to transcriptional activation exists surrounding the transcriptional start site of each gene [1].

The Epstein Barr virus (EBV) has highly controlled and distinct programs of gene expression. The EBV genome consists of ~170 kb of double strand DNA. Following infection of cells the genome is transported to the nucleus, where it becomes associated with histones and so resembles cellular DNA. During viral latency one of the

Janos Minarovits and Hans Helmut Niller (eds.), *Epstein Barr Virus: Methods and Protocols*, Methods in Molecular Biology, vol. 1532, DOI 10.1007/978-1-4939-6655-4_14,

four distinct patterns of viral gene expression is established (termed latency 0 to latency III) [2]. These have relevance for viral persistence and for virus-associated disease. Cells exhibiting these latency patterns can be activated to initiate the viral lytic replication cycle, during which most viral genes are expressed. In addition, an important but currently less well-defined pattern of gene expression occurs immediately following infection of B-cells, and during some diseases, which is termed pre-latency or abortive lytic replication.

The relevant proteins include histones, specific posttranslational modifications of histones, transcription factors, coactivators and corepressors, and proteins that enable and change chromatin architecture. Several EBV genes encode transcription factors that either interact directly (e.g., Zta (Zebra) and Rta), or indirectly (e.g., EBNA1, -2, - 3 and -LP) with DNA [2].

ChIP-Seq has been valuable in determining information about the binding sites of individual EBV transcription factors with the human, and in some cases viral, genomes [3–9]. In addition, the Encyclopedia of DNA elements (ENCODE) project [10] includes an EBV-infected cell displaying latency III pattern of gene expression, and the wealth of data available from ENCODE has been mapped to the EBV genome. This generated an atlas of interactions in Latency III [11].

The ChIP-Seq technique takes a genome-wide approach to determine the interactions of proteins with DNA. This involves the use of cross-linkers to "fix" adjacent protein-protein and protein-DNA interactions within live cells, followed by isolation of a specific protein, together with any other proteins, and DNA that protein is associated with. The precipitated DNA and a sample of the starting DNA are then isolated, the cross-links are reversed and samples are prepared for sequencing. The libraries are then subject to next generation sequencing. The resulting DNA sequencing reads are mapped to genome sequences. For virus-infected cells, this can be undertaken for both the viral genome and the host cell genome from the same sequence library as shown for the viral transcription factor Zta [9, 12].

2 Materials

2.1 Chromatin Preparation and Precipitation

1. Formaldehyde: 37% (v/v).
2. Glycine: 1 M.
3. PBS: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4.
4. Cell Lysis Buffer: 85 mM KCl, 0.5% (v/v) NP-40, and 5 mM PIPES pH 8.0.

5. SDS lysis buffer: 1% (w/v) SDS, 10 mM EDTA, 50 mM Tris pH 8.0.
6. IP Dilution Buffer: 0.01% (w/v) SDS, 1.1% (v/v) Triton X-100, 1.2 mM EDTA, 16.7 mM Tris pH 8.0, 167 mM NaCl.
7. Low salt wash buffer: 0.1% (w/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 150 mM NaCl.
8. High salt wash buffer: 0.1% (w/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 500 mM NaCl.
9. LiCl wash buffer: 250 mM LiCl, 1% (v/v) NP-40, 1% (w/v) Na-deoxycholate, 1 mM EDTA, 10 mM Tris pH 8.0.
10. Elution buffer: 10 mM Tris–HCl, 1 mM EDTA, 1% (w/v) SDS.
11. Protein A sepharose beads.
12. Protein G sepharose beads.
13. Blocking buffer: PBS, 0.5% (w/v) BSA (fraction V), passed through a 0.2 μm filter).
14. RNase A: 20 mg/mL.
15. Proteinase K (20 mg/mL).

2.2 ChIP Library Preparation

1. Library preparation kits: NEB Next ChIP-Seq Library Prep Reagent Set for Illumina, NEBNext Multiplex Oligos for Illumina (Index Primers Set 1).
2. Deoxynucleotide (dNTP) solution mix: Kit component, 10 mM dATP, 10 mM dGTP, 10 mM dCTP, 10 mM dTTP, supplied in Milli-Q water (Millipore Corporation) as a sodium salt at pH 7.5.
3. Phosphorylation reaction buffer (10×): Kit component, 500 mM Tris–HCl, 100 mM MgCl2, 100 mM DTT, 10 mM ATP, pH 7.5 at 25 °C.
4. Enzymes: Kit components, T4 DNA Polymerase, DNA polymerase I: Large (Klenow) fragment, T4 Polynucleotide Kinase, Klenow Fragment (3′→5′exo–), quick T4 DNA ligase, USER enzyme.
5. NEBuffer 2 for Klenow fragment (3′ → 5′exo–): Kit component, 50 mM NaCl, 10 mM Tris–HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9 at 25 °C.
6. Deoxyadenosine 5′- Triphosphate (dATP): Kit component, 1 mM, supplied in Milli-Q water as a sodium salt at pH 7.5.
7. Quick ligation reaction buffer (2×): Kit component, 132 mM Tris–HCl, 20 mM $MgCl_2$, 2 mM dithiothreitol, 2 mM ATP, 15% (v/v) Polyethylene glycol (PEG 6000), pH 7.6 at 25 °C.
8. NEBNext Adaptor: Kit component, 15 μM, 5′-/5Phos/GAT CGG AAG AGC ACA CGT CTG AAC TCC AGT C/

ideoxyU/A CAC TCT TTC CCT ACA CGA CGC TCT TCC GAT C*T-3′, where the * indicates a phosphorothioate bond.

9. High-Fidelity 2× PCR Master Mix: Sample kit component.
10. Universal PCR primer: Kit component, 25 μM (*see* **Note 1**), 5′-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC*T-3′, where the * indicates a phosphorothioate bond.
11. NEBNext Multiplex Oligos for Illumina: Kit component, e.g., Index 1: 5′-CAA GCA GAA GAC GGC ATA CGA GAT CGT GAT GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC-s-T-3′, where -s- indicates a phosphorothioate bond (*see* **Note 2**).
12. Qiagen MinElute PCR Purification and Gel Extraction Kits.
13. PB buffer: Kit component, high concentration of guanidine hydrochloride and isopropanol.
14. MinElute Columns: Kit component.
15. QIAquick spin columns: Kit component.
16. 2 mL collection tubes: Kit component.
17. PE buffer: Kit component.
18. EB buffer: Kit component, 10 mM Tris-Cl, pH 8.5.
19. QG buffer: Kit component.
20. Thermal cycler.
21. Agarose.
22. TAE buffer (1×): 4.84 g Tris base, 1.14 mL acetic acid, pH 8.0. Make up to 1 L with sterile water.
23. Loading buffer: 50 mM Tris, pH 8.0, 40 mM EDTA, 40 % w/v sucrose.
24. 100 bp DNA ladder (New England Bio Labs Quick-Load 100 bp DNA Ladder).
25. 1 kb DNA ladder (Bioline HyperLadder).
26. Electrophoresis equipment (Bio-Rad Mini-Sub Cell GT System #170-4467 and power source).
27. Dark Reader Transilluminator (Clare Chemical Research).
28. Disposable scalpels.
29. 96–100 % (v/v) ethanol.
30. Isopropanol.
31. Microcentrifuge.
32. Agilent 2100 Bioanalyzer.
33. Agilent High Sensitivity DNA Kit.
34. High Sensitivity DNA Chip: Kit component.

35. Electrode Cleaner: Kit component.
36. Syringe: Kit component.
37. Spin Filters: Kit component.
38. High Sensitivity DNA Ladder: Kit component, 13 DNA fragments from 50 to 7000 bp.
39. High Sensitivity DNA Markers: Kit component, lower marker at 35 bp, upper marker at 10,380 bp.
40. High Sensitivity DNA Dye Concentrate: Kit component.
41. High Sensitivity DNA Gel Matrix: Kit component.
42. IKA MS3 vortex mixer with chip adaptor (Agilent).
43. Electrode cleaner (Agilent).

3 Methods

3.1 Chromatin Preparation and Precipitation

1. Centrifuge 1×10^8 cells at $1710 \times g$ for 5 min, resuspend in 5 mL of cell culture medium to 2×10^7/mL.
2. Add formaldehyde to cell culture medium to a final concentration of 1% and incubate for 15 min at 20 °C while rocking.
3. Add 625 μL of 1 M glycine, to a final concentration of 0.125 M.
4. Centrifuge cells at $1710 \times g$ for 5 min resuspend in 5 mL of PBS.
5. Centrifuge cells at $1710 \times g$ for 5 min resuspend cells in 3 mL of Cell Lysis Buffer.
6. Incubate on ice for 10 min.
7. Centrifuge at $1710 \times g$ 5 min, and discard supernatant (*see* **Note 3**).
8. Resuspend nuclei in 2 mL SDS lysis buffer.
9. Sonicate nuclear lysate on ice (e.g., using a Branson Model 250 Microtip sonicator) 10×10 s pulses, in a volume of 100 μL with a 21% output, vortexing the sample after every minute (*see* **Note 4**).
10. Snap freeze and store nuclear chromatin extracts at −80 °C.
11. Mix 1 mL of Protein-A Sepharose beads with 1 mL of Protein-G Sepharose beads (*see* **Note 5**) and 10 mL of IP Dilution Buffer (*see* **Note 6**).
12. Centrifuge beads at $1710 \times g$ for 5 min, and discard supernatant.
13. Add 10 mL of IP Dilution Buffer, centrifuge beads at $1710 \times g$ for 5 min and discard supernatant.
14. Resuspend beads in 5 mL of blocking buffer and incubate at 4 °C while rotating for 30 min.

15. Centrifuge beads at 1710×*g* for 5 min and discard supernatant.
16. Resuspend beads in 2 mL of IP dilution buffer.
17. Defrost 2 mL of chromatin and transfer it to a 50 mL falcon tube.
18. Add 18 mL IP dilution buffer.
19. Remove 400 μL and freeze. Label as "input chromatin."
20. Add 1 mL of the blocked protein A-G beads to the remaining 19.4 mL and incubate for 30 min at 4 °C.
21. Centrifuge beads at 1710×*g* for 5 min and transfer supernatant into a new 50 mL tube.
22. Add 40 μg of antibody and incubate for 1 h on a rotator.
23. Add 1 mL of blocked protein A-G beads and rotate tubes overnight.
24. Centrifuge beads at 1710×*g* for 5 min and add 20 mL low salt wash buffer. Incubate for 10 min on a rotator.
25. Centrifuge beads at 1710×*g* for 5 min and add 20 mL high salt wash buffer. Incubate for 10 min on a rotator.
26. Centrifuge beads at 1710×*g* for 5 min and add 20 mL LiCl wash buffer. Incubate for 10 min on a rotator.
27. Centrifuge beads at 1710×*g* for 5 min and add 20 mL TE buffer. Incubate for 10 min on a rotator.
28. Centrifuge beads at 1710×*g* for 5 min and discard supernatant.
29. Add 3 mL of TE/1% (v/v) SDS at room temperature and incubate the solution at 65 °C for 20 min.
30. Centrifuge beads at 1710×*g* for 5 min. Transfer the supernatant to a new tube.
31. Defrost the "input chromatin" control and add 3 mL of TE/1% (w/v) SDS at room temperature.
32. Incubate the input chromatin and the precipitated chromatin at 65 °C overnight to reverse the cross-links.
33. Add 60 μL of RNAse A and 3 mL of TE to each and incubate at 37 °C for 2 h (*see* **Note 7**).
34. Add 60 μL of proteinase K to each tube and incubate at 55 °C for 2 h (*see* **Note 7**).
35. Purify DNA using the Qiagen MinElute PCR purification kit (*see* **Note 11**). Add ethanol (96–100%) to buffer PE before use. All centrifugation steps are carried out in a microcentrifuge set to 11,710×*g* at room temperature (*see* **Note 12**).
36. Add 15 mL PB buffer to the samples and mix.
37. Place the MinElute column in a provided 2 mL collection tube.

38. Apply 700 μL of the samples to the column and centrifuge for 1 min.
39. Discard the flow-through and place the MinElute column back into the same tube.
40. Repeat **steps 38** and **39** until all of the sample has been centrifuged.
41. Add 750 μL PE buffer to the column, let it stand for 3 min, and centrifuge for 1 min.
42. Discard the flow-through and place the column back into the same tube. Centrifuge the column for an additional minute.
43. Place the MinElute column into a sterile 1.5 mL microcentrifuge tube.
44. Add 42 μL EB buffer to the center of the membrane, let the column stand for 1 min, then centrifuge for 1 min.
45. Add the eluate back to the center of the column, let it stand for 1 min, then centrifuge for 1 min.

3.2 ChIP Library Preparation

1. The NEBNext ChIP-Seq Library Prep Reagent Set is used in combination with the NEBNext Multiplex Oligos for Illumina (Index Primers Set 1), the MinElute PCR purification kit (Qiagen), and the Gel Extraction kit (Qiagen) to prepare the input and ChIP libraries.
2. To end-repair the ChIP and input DNA, measure the concentration of your ChIP and input samples (*see* **Note 8**).
3. Suspend 10 ng (*see* **Note 9**) of ChIP or input DNA in sterile water to a final volume of 40 μL.
4. Dilute DNA polymerase I, Large (Klenow) Fragment by mixing 1 μL of enzyme with 4 μL of sterile water.
5. Mix the following components in a sterile microcentrifuge tube:
   (a) 20 μL of Phosphorylation Reaction Buffer (10×).
   (b) 4 μL of T4 DNA Polymerase.
   (c) 4 μL of T4 polynucleotide kinase.
   (d) 8 μL of dNTP mix.
   (e) 4 μL of diluted DNA Polymerase I, Klenow Fragment.
6. Add 10 μL of this mix to each sample of the 40 μL ChIP and input DNA.
7. Incubate in a thermal cycler at 20 °C for 30 min (*see* **Note 10**).
8. Purify DNA using the Qiagen MinElute PCR purification kit (*see* **Note 11**). Add ethanol (96–100% (v/v)) to buffer PE before use. All centrifugation steps are carried out in a micro-

centrifuge set to 11,710×*g* at room temperature (*see* **Note 12**).

9. Add 250 μL PB buffer to the samples and mix.
10. Place the MinElute column in a provided 2 mL collection tube.
11. Apply all traces of the samples to the column and centrifuge for 1 min.
12. Discard the flow-through and place the MinElute column back into the same tube.
13. Add 750 μL PE buffer to the column, stand for 3 min, and centrifuge for 1 min.
14. Discard the flow-through and place the column back into the same tube. Centrifuge the column for an additional minute.
15. Place the MinElute column into a sterile 1.5 mL microcentrifuge tube.
16. Add 35 μL EB buffer to the center of the membrane, let the column stand for 1 min, then centrifuge for 1 min.
17. Add the eluate back to the center of the column, let it stand for 1 min, then centrifuge for 1 min.
18. For dA-tailing of end-repaired DNA, mix the following components in a sterile microcentrifuge tube:
    (a) 20 μL NEBuffer 2 (10×).
    (b) 40 μL Deoxyadenosine 5′-Triphosphate.
    (c) 4 μL Klenow Fragment (3′→5′ exo–).
19. Add 16 μL of this mix to the 34 μL of DNA sample (*see* **Note 13**).
20. Incubate in a thermal cycler for 30 min at 37 °C.
21. Repeat **steps 8–17**, but using 11 μL EB buffer.
22. For the adaptor ligation, dilute the NEBNext Adaptor for Illumina (15 μM) 10-fold in sterile water to a final concentration of 1.5 μM.
23. Mix the following components in a sterile microcentrifuge tube:
    (a) 60 μL of Quick Ligation Reaction Buffer (2×).
    (b) 4 μL of Diluted NEBNext Adaptor (1.5 μM).
    (c) 16 μL of Quick T4 ligase.
24. Add 20 μL of this mix to the 10 μL end-repaired, dA-tailed DNA.
25. Incubate the reaction in a thermal cycler for 15 min at 20 °C (*see* **Note 10**).
26. Add 3 μL USER enzyme, mix by pipetting up and down, and incubate at 37 °C for 15 min.

27. Repeat **steps 8–17**, but using 150 μL PB and 24 μL EB buffer.
28. To PCR enrich the adaptor ligated DNA, mix the following components in a sterile microcentrifuge tube:
    (a) 23 μL adaptor ligated DNA.
    (b) 25 μL NEBNext High-Fidelity 2× PCR Master Mix.
    (c) 1 μL universal PCR primer (25 μM).
    (d) 1 μL Index Primer (one of 12 per sample, e.g., Input: Index 4, ChIP1: Index 6, ChIP2: Index 12).
29. PCR cycling conditions (Table 1):
30. Repeat **steps 8–17**, but using 16 μL EB buffer.
31. To size select the library, prepare the required number of 2% (w/v) agarose gels with 1× TAE buffer (*see* **Note 14**).
32. Add 6 μL loading buffer to the samples and to 15 μL of the 100 bp DNA ladder.
33. Use 1× TAE buffer as running buffer. Load the sample and marker ladders onto the gel(s), leaving at least one space between ladder and samples (*see* **Note 15**).
34. Run the gel at 120 V until the marker dye has reached the bottom of the gel (approximately 60 min).
35. Weigh out three sterile microcentrifuge tubes and label them with their weight.
36. View the gels on the Dark Reader Transilluminator (*see* **Note 16**).
37. Excise the 175–225 bp region of the gel with a clean, disposable scalpel (*see* **Note 17**). Minimize the size of the gel slice by removing excess agarose. Place the selected gel slice into a labeled microcentrifuge tube.
38. Use the Qiagen gel extraction kit to purify the DNA from the agarose slice. All centrifugation steps are carried out in a microcentrifuge set to 11,710 × *g* at room temperature (*see* **Note 12**).

Table 1
PCR cycling conditions

Cycle step	Temperature (°C)	Time (s)	Cycles
Initial denaturation	98	30	1
Denaturation	98	10	15
Annealing	65	30	
Extension	72	30	
Final extension	4	hold	1

39. Weigh the tubes containing the gel slice and calculate the weight of the agarose.
40. Add three volumes of QG buffer to one volume of gel (i.e., add 300 μL of QG to 100 mg of gel). For gel slices over 400 mg, use more than one tube.
41. Incubate on bench (*see* **Note 18**) for at least 10 min until the gel slice has completely dissolved. Vortex the tube every 2 min during the incubation time.
42. Heat EB buffer to 50 °C.
43. Add one gel volume of isopropanol to the sample (i.e., if the gel slice weight 100 mg, add 100 μL) and mix (*see* **Note 19**).
44. Place the QIAquick spin column in a provided 2 mL collection tube.
45. Apply the sample to the column and centrifuge for 1 min. The maximum volume of the column is 750 μL. For higher sample volumes, load the spin column again after centrifugation.
46. Discard flow-through and place the column back into the same collection tube.
47. Add 500 μL QG buffer to the column and centrifuge for 1 min.
48. Discard flow-through.
49. Add 750 μL of PE buffer to the column and let it stand for 3 min, then centrifuge for 1 min.
50. Discard flow-through and centrifuge the column for an additional minute.
51. Place the column into a sterile 1.5 mL microcentrifuge tube.
52. Add 55 μL of EB buffer directly to the center of the column membrane and let the column stand for 1 min, then centrifuge for 1 min.
53. Repeat **steps 8–17**, but using 14 μL EB buffer.
54. Use 1 μL of the purified library sample for quality control on a Bioanalyzer (*see* **Note 20**) using the Agilent High Sensitivity DNA Kit:
55. Let the High Sensitivity DNA dye concentrate and High Sensitivity DNA gel matrix equilibrate to room temperature (*see* **Note 21**).
56. Vortex the vial with High Sensitivity DNA dye concentrate for 10 s and spin down. Making sure that the DMSO is completely thawed.
57. Pipette 15 μL of the dye concentrate into a High Sensitivity DNA gel matrix vial (*see* **Note 22**).
58. Return the dye concentrate to 4 °C to store, protecting it from light.

59. Cap the tube, vortex for 10 s. Make sure that the gel and dye are mixed properly.
60. Transfer the gel—dye mix to the top receptacle of the spin filter.
61. Spin for 10 min at 2400×*g* at room temperature in a microcentrifuge.
62. Discard the filter and label the tube (*see* **Note 23**).
63. Ensure that the base plate of the chip priming station is in position C and that the adjustable clip is set to the lowest position.
64. Place a new High Sensitivity DNA chip onto the chip priming station.
65. Placing the pipette tip at the center of the well, pipette 9 μL of the gel—dye mix (which is at room temperature) to the bottom of the well located in the right (fourth) column and third row.
66. Set a timer to 1 min and make sure that the plunger is positioned at 1 mL. Then close the chip priming station.
67. Press the syringe plunger down until it is held by the clip, then start the timer.
68. Release the plunger after exactly 1 min.
69. Make sure that the plunger moves back at least to the 0.3 mL mark.
70. Wait 5 s before slowly pulling back the plunger to the 1 mL position.
71. Open the chip priming station and pipette 9 μL of the gel—dye mix into the three empty wells of the right (fourth) column.
72. Return the gel—dye mix to 4 °C to store, protecting it from light.
73. Pipette 5 μL of the High Sensitivity DNA marker into each of the empty wells (first three columns) (*see* **Note 24**).
74. Pipette 1 μL of the High Sensitivity DNA ladder into the well marked with the ladder symbol (third column, fourth row).
75. Pipette 1 μL of sample (*see* **Note 25**) or 1 μL of marker (for otherwise empty wells) into each well of the first three columns, except the well with the ladder symbol.
76. Carefully place the chip horizontally into the vortex adapter and vortex for 1 min at 2400 rpm.
77. Place the chip carefully into the receptacle of the Bioanalyzer, making sure that the run is started within 5 min.
78. Carefully close the lid (*see* **Note 26**) and start the chip run. Use the dsDNA assay and fill in the sample name table.
79. After the run is finished dispose of the chip.
80. Slowly fill one of the wells of the electrode cleaner (*see* **Note 27**) with 350 μL sterile water.

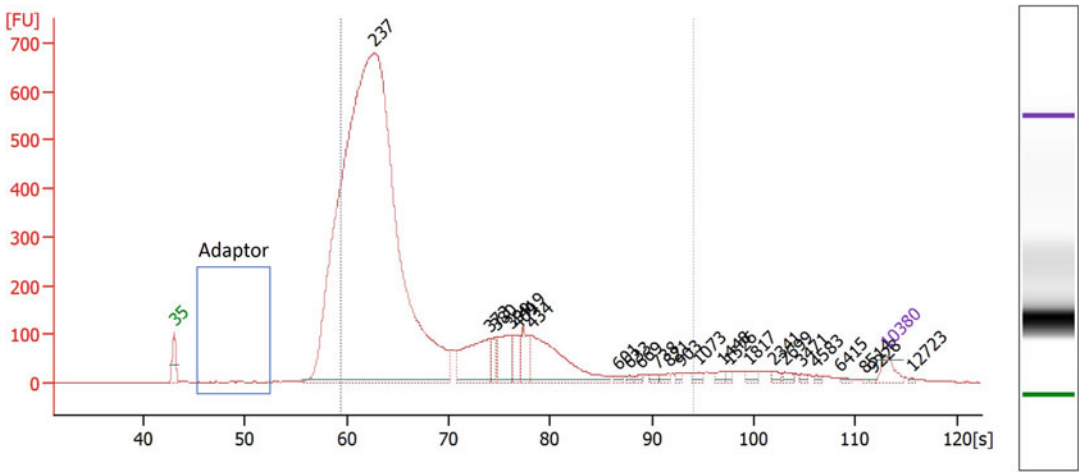


Fig. 1 Bioanalyzer Electropherogram showing fluorecense units [FU] on the *y*-axis and seconds [s] on the *x*-axis, with numbers above peaks indicating the fragment size in base pairs (bp). A single high peak around 200 bp and no peak in the 100 bp Adapotor region has been detected

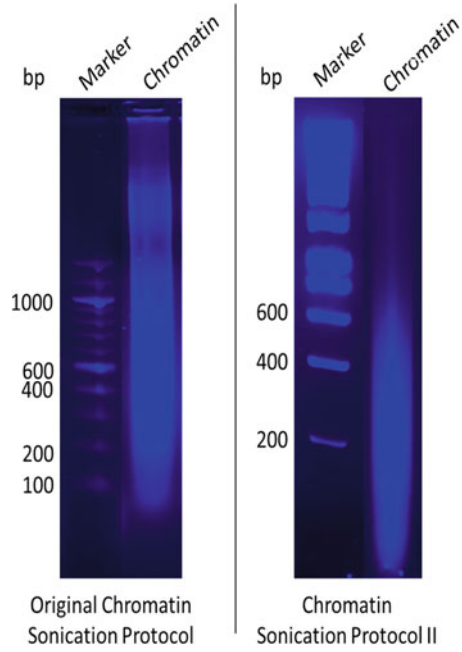


Fig. 2 Agarose gel showing the size, in base pairs (bp), of sonicated chromatin fragments, using two different chromatin sonication protocols (*left* and *right*)

81. Place the electrode cleaner in the Bioanalyzer, then close the lid and leave it for about 10 s.
82. Remove the electrode cleaner and wait another 10 s to allow the water on the electrodes to evaporate.
83. For a good quality sample, the electropherogram should show no peak in the 100 bp region, with a high peak around 200 bp (*see* Fig. 1). If there is a peak in the adaptor region (as indicated by the blue box), the sample is contaminated and sequencing may not yield successful results.

4 Notes

1. Concentrations of universal PCR Primer and the Index Primers contained in the NEB kit have changed from 25 to 10 μM. Refer to the most recent protocol for use.
2. 12 Index Primers are included in the NEBNext Multiplex Oligos for Illumina (Index Primers Set 1) kit for producing barcoded libraries. Refer to the manual for sequence information. If fewer than 2 indexes per lane are used, NEB recommends the combination of the following indexes for best demultiplexing results:
   (a) Pool of 2 samples: Index 6 and 12.
   (b) Pool of 3 samples: Index 4, 6, and 12.
   (c) Pool of 6 samples: Index 2, 4, 5, 6, 7, and 12.
3. After the addition of cell lysis buffer all steps are undertaken at 4 °C unless specified.
4. Sonication is critical. The ideal fragment size for ChIP sequencing is 200 bp. Appropriate conditions will vary with each sonicator and can be checked by undertaking a time course of sonication, incubating at 65 °C overnight to reverse the cross-links. This is followed by purifying the DNA and analyzing size distribution on an agarose gel. Figure 2 shows the results of two different sonication protocols. The gel on the left shows a sub-ideal chromatin sonication for ChIP-Seq purposes. The gel on the right, using the second chromatin sonication protocol, shows a good fragmentation around 200 bp.
5. Protein A and Protein G sepharose beads should be stored at 4 °C in 70% (v/v) ethanol, then washed with IP dilution buffer prior to use.
6. Add protease inhibitor cocktail (Sigma) and 1 mM PMSF to all buffers just before use.
7. Proteinase K and RNAase can be stored as aliquots at −20 °C.

8. We found that the Qubit Fluorometer worked well for this step. The concentration of our ChIP samples was too low to be measured accurately using a NanoDrop Spectrophotometer.
9. The kit calls for 10 ng of DNA, but we have previously used 18 ng to ensure maximum library concentration. In our experience, this did not seem to impair the quality of the final library.
10. Our laboratory tends to be around 20 °C. We incubate this reaction on the bench rather than in a thermal block cycler and still get good results.
11. If using the Qiagen kit, it is important for the efficiency of this purification process that centrifugation steps are actually carried out at room temperature.

    Alternatively, it is possible to clean up the sample using AMPure XP Beads (Beckman Coulter, Inc.).
12. Qiagen suggests that buffer QG from the GelExtraction kit may be used to remove salt and proteins from samples. Do not use buffer QG outside of the gel extraction during the library preparation.
13. The sample was eluted in 35 μL, but a small amount will remain in the column; hence, the DNA is already in the correct 34 μL volume required for the following step.
14. Make sure that the comb creating the sample wells in the agarose gel is precisely straight to ensure that the sample and markers run exactly vertically on the gel. If the samples and markers run at an angle, it is more difficult to be sure of the exact size to select on the gel, which can lead to adaptor contamination.
15. It is important to select the correct sample size, we suggest running each sample with two markers on each side, as shown in the scheme below.

Marker 1	Marker 2	Sample	Marker 1	Marker 2

16. The Dark Reader Transilluminator allows the visualization of the DNA on the gel for excision of the desired size while avoiding UV exposure that may be harmful to the DNA sample.
17. Make sure not to cut out any of the light adaptor band (below 150 bp). It is advisable to take a photograph of the gel before and after the excision.
18. Do not heat the samples to 55 °C as stated in the GelExtraction kit protocol.
19. Do not centrifuge the sample at this point.
20. If you are unable to perform quality control of your sample using a bioanalyzer, your sequencing facility may be able to perform this step for you. In this case, we recommend per-

forming a qPCR analysis of a target that is known/thought to be enriched before sending your samples off.

21. Make sure to protect the DNA dye concentrate from light.
22. Use the indicated volumes. Different volumes at the same ratio can produce inaccurate results.
23. Shield the gel—dye mix from light and store at 4 °C. It is sufficient for five chips. Use within 6 weeks of preparation.
24. Even if you have less than 11 samples, do not leave any wells empty; otherwise, the chip will not run properly.
25. Samples should be dissolved in 10 mM Tris and 1 mM EDTA for optimal results. However, we have achieved good results with our samples in EB buffer.
26. Never use force to close the lid or drop the lid onto the chip. This may damage the electrodes or cause liquid spills leading to bad results.
27. A more thorough clean may be required when switching between different types of assay.

Acknowledgment

This work was supported by MRC MR/J001 708/1.

References

1. Kimura H (2013) Histone modifications for human epigenome analysis. J Hum Genet 58:439–445
2. Longnecker R, Kieff E, Cohen JI (2013) In: Knipe D, Howley P (eds.), Fields Virology, 6th edn. Lippincott Williams & Wilkins (Philadelphia, Pennsylvania, United States)
3. Bergbauer M, Kalla M, Schmeinck A, Gobel C, Rothbauer U, Eck S, Benet-Pages A, Strom TM, Hammerschmidt W (2010) CpG-methylation regulates a class of Epstein-Barr virus promoters. PLoS Pathog 6:e1001114
4. Holdorf MM, Cooper SB, Yamamoto KR, Miranda JL (2011) Occupancy of chromatin organizers in the Epstein-Barr virus genome. Virology 415:1–5
5. Ramasubramanyan S, Osborn K, Flower K, Sinclair AJ (2012) Dynamic chromatin environment of key lytic cycle regulatory regions of the Epstein-Barr virus genome. J Virol 86:1809–1819
6. Woellmer A, Arteaga-Salas JM, Hammerschmidt W (2012) BZLF1 governs CpG-methylated chromatin of Epstein-Barr Virus reversing epigenetic repression. PLoS Pathog 8:e1002902
7. Jiang S, Willox B, Zhou H, Holthaus AM, Wang A, Shi TT, Maruo S, Kharchenko PV, Johannsen EC, Kieff E et al (2014) Epstein-Barr virus nuclear antigen 3C binds to BATF/IRF4 or SPI1/IRF4 composite sites and recruits Sin3A to repress CDKN2A. Proc Natl Acad Sci U S A 111:421–426
8. McClellan MJ, Wood CD, Ojeniyi O, Cooper TJ, Kanhere A, Arvey A, Webb HM, Palermo RD, Harth-Hertle ML, Kempkes B et al (2013) Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. PLoS Pathog 9:e1003636
9. Ramasubramanyan S, Osborn K, Al-Mohammad R, Naranjo Perez-Fernandez IB, Zuo J, Balan N, Godfrey A, Patel H, Peters G, Rowe M et al (2015) Epstein-Barr virus transcription factor Zta acts through distal regulatory elements to directly control cellular gene expression. Nucleic Acids Res 43:3563–3577

10. Kellis M, Wold B, Snyder MP, Bernstein BE, Kundaje A, Marinov GK, Ward LD, Birney E, Crawford GE, Dekker J et al (2014) Defining functional DNA elements in the human genome. Proc Natl Acad Sci U S A 111:6131–6138
11. Arvey A, Tempera I, Tsai K, Chen HS, Tikhmyanova N, Klichinsky M, Leslie C, Lieberman PM (2012) An atlas of the Epstein-Barr virus transcriptome and epigenome reveals host-virus regulatory interactions. Cell Host Microbe 12:233–245
12. Ramasubramanyan S, Kanhere A, Osborn K, Flower K, Jenner RG, Sinclair AJ (2012) Genome-wide analyses of Zta binding to the Epstein-Barr virus genome reveals interactions in both early and late lytic cycles and an epigenetic switch leading to an altered binding profile. J Virol 86:12494–12502